MEMOIRS
OF THE
RAMAN RESEARCH INSTITUTE

Vol. VII

(Nos. 110 to 120)

BANGALORE

1959

4 JAN 1965
MADRAS

MEMOIRS OF THE RAMAN RESEARCH INSTITUTE

Vol. VII, Nos. 110 *to* 120

110.	Light propagation in absorbing crystals possessing optical activity—Electromagnetic theory ..	S. PANCHARATNAM
111.	Percussion figures in crystals	C. V. RAMAN
112.	The Raman spectra of crystalline sulphates of Ni and Mn	D. KRISHNAMURTI
113.	The relativistic theory of chemical binding ..	K. S. VISWANATHAN
114.	Christiaan Huyghens and the wave theory of light	C. V. RAMAN
115.	The optics of mirages	C. V. RAMAN AND S. PANCHARATNAM
116.	Caustics formed by diffraction and the geometric theory of diffraction patterns	C. V. RAMAN
117.	Huyghens' principle and the diffraction of light. Part I. Theoretical considerations	C. V. RAMAN
118.	The optics of mirages	C. V. RAMAN
119.	The Raman and infra-red spectra of some solid hydroxides. Parts I, II and III	D. KRISHNAMURTI
120.	X-ray study of the structure of moonstones ..	A. JAYARAMAN

Reprinted from "The Proceedings of the Indian Academy of Sciences",
Vol. XLVIII, No. 4, Sec. A, 1958

4 JAN

LIGHT PROPAGATION IN ABSORBING CRYSTALS POSSESSING OPTICAL ACTIVITY—ELECTROMAGNETIC THEORY

BY S. PANCHARATNAM

(*Memoir No.* 110 *from the Raman Research Institute, Bangalore*-6)

Received September 24, 1958
(Communicated by Sir C. V. Raman)

§ 1. INTRODUCTION

IN the present paper we shall give the electromagnetic theory of light propagation in absorbing crystals possessing optical activity: the theoretical presentation is a straightforward extension of that previously adopted for transparent active crystals.[1] We shall find that the results of the electromagnetic theory are, for most practical purposes, the same as obtained previously[2] on the basis of a simpler physical idea, *viz.*, by a 'method of superposition'—the results of which have been confirmed by detailed observations on the interesting phenomena displayed by amethyst.[3]

A complete solution of the electromagnetic theory of light propagation in absorbing active crystals has thus far been presented only for the case of uniaxial media.[4] Recently an entirely new and general method of formally solving the electromagnetic equations has been introduced by Jones[5] in a paper dealing with light propagation in anisotropic media (*see*, however, Appendix). We shall find in § 11, that it becomes essential to adopt this new method for solving the propagation of light along certain remarkable directions, *viz.*, the so-called *singular axes* which can exist both in active[3] as well as inactive[6,7] crystals.

§ 2. FORMULATION OF THE PROBLEM

Consider an arbitrary direction of propagation Oz in the crystal—which direction we may conveniently take as being normal to the plane of the paper. Our problem is merely to determine the characteristic states of polarisation of the plane waves that can be propagated along this direction, as well as their velocities and absorption coefficients. The reason why only specific types of waves can be propagated along the z-direction is that the field vectors (in particular, the vectors **D** and **E**) of the electromagnetic wave are constrained to satisfy certain relations amongst them-

selves. Firstly they must satisfy Maxwell's equations; and secondly they must obey certain constitutive relations imposed by the properties of the medium, and which determine the optical characteristics of the medium. These latter relations will be determined by the polarisable characteristics of the medium, *i.e.*, the relation which the induced polarisation bears to the electric field of the light wave. (As is customary even in the case of transparent active crystals we shall assume for simplicity that there is no induced magnetization, *i.e.*, that $\mathbf{B} = \mathbf{H}$, though—as in that case—this assumption appears to violate energy considerations.)

§ 3. MAXWELL'S RELATIONS FOR A PLANE WAVE FIELD

Since we are interested in plane waves propagating in the z-direction, the field vectors do not vary over planes normal to the z-direction. If $\mathbf{k}$ denote a unit vector along the z-axis, we may use the operator $\mathbf{k}\,(\partial/\partial z)$ in place of the gradient operator ∇ in the Maxwell's equations—or rather in the standard relation obtained by eliminating $\mathbf{H}$ between the Maxwell's equations. The latter relations (*see*, *e.g.*, Page and Adams,[8] eq. 82–9) then assume the form

$$\ddot{\mathbf{D}} = c^2 \frac{\partial^2}{\partial z^2} [\mathbf{E} - \mathbf{k}\,(\mathbf{k} \cdot \mathbf{E})]. \tag{1}$$

The components of this equation take the simpler form

$$\ddot{\mathrm{D}}_x = c^2 \frac{\partial^2 \mathrm{E}_x}{\partial z^2}\,; \qquad \ddot{\mathrm{D}}_y = c^2 \frac{\partial^2 \mathrm{E}_y}{\partial z^2} \tag{2}$$

$$\ddot{\mathrm{D}}_z = 0. \tag{3}$$

The last condition implies the transversality of the displacement vector—since we shall be concerned only with fields varying harmonically in time. Also we shall first restrict our consideration to homogeneously polarised plane waves, so that

$$\mathbf{D}, \mathbf{E} \sim \exp .\, i\omega \left(t - \frac{z}{c} \cdot \bar{n}\right) \times \mathbf{D}_0,\ \mathbf{E}_0 \tag{4}$$

where $\bar{n}$ is the complex refractive index. Under these conditions the operation $\partial/\partial z$ becomes identical with multiplication by $-i\omega\bar{n}/c$ and the operation $\partial/\partial t$ with multiplication by $i\omega$. Thus for a plane wave field of type (4) Maxwell's equations (with $\mathbf{H}$ eliminated) finally assume the elegant form

$$\bar{v}^2 \mathrm{D}_x = c^2 \mathrm{E}_x\,; \qquad \bar{v}^2 \mathrm{D}_y = c^2 \mathrm{E}_y \tag{5}$$

where $\bar{v}$ represents the complex velocity $c/\bar{n}$ of the damped wave.

§ 4. Properties of the Medium

The dielectric displacement in the medium depends on the electric vector, and, for a transparent optically active crystal,[9] **D** may be expressed as an explicit vector function of **E** by the use of the dielectric tensor (ϵ) and gyration tensor (g); it was shown in a previous paper[1] that it was more convenient to reverse the procedure and express **E** as an explicit function of **D** by the use of the index-tensor (a), and a modified gyration tensor (γ). When we turn to media possessing absorption, the components of the dielectric tensor and the gyration tensor will become complex quantities: in terms of our modified presentation it may be shown that the consequence of this is that the constants of the index tensor and the modified gyration tensor have to be replaced by complex quantities (*see* Appendix). Thus the relation between **E** and **D** may be expressed in the same form as previously given for transparent active crystals (reference 1, eq. 2 *a*).

$$c^2\mathbf{E} = (\bar{a})\,\mathbf{D} - i\bar{\Gamma} \times \mathbf{D}. \tag{6}$$

Here $(\bar{a})$ is a symmetric tensor—the complex index-tensor; and $\bar{\Gamma}$ is an auxiliary vector which depends on the direction of the wave-normal **s**, being given by

$$\bar{\Gamma} = (\bar{\gamma})\,\mathbf{s} \tag{7}$$

where $(\bar{\gamma})$ represents the 'tensor of optical activity'—a general tensor with complex components. It is convenient to express the above relations in terms of tensors having real components; *these tensors will in turn separately determine the various optical characteristics of the medium.*

Thus in (6) and (7) we may substitute

$$(\bar{a}) = (a) + i\,(b) \tag{8}$$

$$(\gamma) = (\gamma) + i\,(\gamma'). \tag{9}$$

Here (a) and (b) are the usual index- and absorption-tensors which occur for example in optically inactive absorbing crystals, and which define the *index-* and *absorption-ellipsoids*; (γ) is the tensor of optical rotation which was referred to in our paper on transparent active crystals[1] as the modified gyration tensor. The new tensor (γ') may be called the tensor of circular

dichroism for reasons which will become apparent as we proceed. In eq. (6) it is clear that we may further substitute

$$\bar{\Gamma} = \Gamma + i\Gamma' \tag{10}$$

where

$$\Gamma = (\gamma)\ \mathrm{s} \tag{11}$$

$$\Gamma' = (\gamma')\ \mathrm{s}. \tag{12}$$

§ 5. SOLUTION OF THE ELECTROMAGNETIC EQUATIONS

We now write down the x and y components of the vector equation (6) after expanding the cross-product and omitting terms in D_z (since the latter is equal to zero).

$$\left.\begin{aligned} c^2 E_x &= \bar{a}_{11} D_x + (\bar{a}_{12} + i\bar{\Gamma}_z)\, D_y \\ c^2 E_y &= (\bar{a}_{12} - i\bar{\Gamma}_z)\, D_x + \bar{a}_{22} D_y \end{aligned}\right\} \tag{13}$$

Substituting for c^2E_x and c^2E_y from (5) we obtain as our fundamental equations:

$$\left.\begin{aligned} \bar{v}^2 - \bar{a}_{11} &= (\bar{a}_{12} + i\bar{\Gamma}_z)\left(\frac{D_y}{D_x}\right) \\ \bar{v}^2 - \bar{a}_{22} &= (\bar{a}_{12} - i\bar{\Gamma}_z)\left(\frac{D_x}{D_y}\right) \end{aligned}\right\} \tag{14}$$

In general there will be two pairs of roots $\bar{v}_a$, $(D_y/D_x)_a$ and $\bar{v}_b$ $(D_y/D_x)_b$ which will satisfy these simultaneous equations. These will give the complex velocities and states of polarisation of the two waves that can be propagated in the z-direction. If we multiply the two equations of (14) to eliminate (D_y/D_x) we obtain the following quadratic in $\bar{v}^2$ whose roots $\bar{v}_a^2$ and $\bar{v}_b^2$ determine the complex velocities (*i.e.*, the velocities and absorption coefficients of the waves):

$$(\bar{v}^2 - \bar{a}_{11})(\bar{v}^2 - \bar{a}_{22}) = \bar{a}_{12}^2 + \bar{\Gamma}_z^2. \tag{15}$$

Subtracting the second equation of (14) from the first to eliminate $\bar{v}^2$, we get

$$(\bar{a}_{12} + i\bar{\Gamma}_z)\left(\frac{D_y}{D_x}\right) - (\bar{a}_{12} - i\bar{\Gamma}_z)\left(\frac{D_x}{D_y}\right) = -(\bar{a}_{11} - \bar{a}_{22}).$$

Or,

$$(\bar{a}_{12} + i\bar{\Gamma}_z)\left(\frac{D_y}{D_x}\right)^2 + (\bar{a}_{11} - \bar{a}_{22})\left(\frac{D_y}{D_x}\right) - (\bar{a}_{12} - i\bar{\Gamma}_z) = 0. \tag{16}$$

This is a quadratic in (D_y/D_x) and *its roots specify the states of polarisation of the waves.* Since these roots will in general be complex, the waves will be elliptically polarised.

The task of discussing in greater detail the velocities, absorption coefficients and states of polarisation of the waves is complicated by the fact that all the coefficients occurring in (15) and (16) are really complex quantities:

$$\bar{a}_{hk} = a_{hk} + ib_{hk}\ ; \quad \bar{\Gamma}_z = \Gamma_z + i\Gamma_z'. \tag{17}$$

It is worthwhile pointing out that the a_{hk} and b_{hk} (which are the components of the index and absorption tensors respectively) are also the constants occurring in the equation to the elliptic sections of the index- and absorption-ellipsoids by the xy plane; in other words the equation to these elliptic sections are respectively

$$\left.\begin{aligned} a_{11}x^2 + a_{22}y^2 + 2a_{12}xy = 1 \\ b_{11}x^2 + b_{22}y^2 + 2b_{12}xy = 1 \end{aligned}\right\}. \tag{18}$$

Also Γ_z is the scalar parameter of optical rotation already met with in the case of transparent optically active crystals;[1] whilst Γ_z' may be called the scalar parameter of circular dichroism for reasons which will be discussed in the next section.

§ 6. Circular Dichroism and Its Directional Variation

The characteristic effect introduced by the presence of the parameter Γ_z' may be best revealed by supposing linear birefringence and linear dichroism to be absent, *i.e.*, by setting [in equations (15) and (16)],

$$a_{11} = a_{22} = a; \quad a_{12} = 0; \quad b_{11} = b_{22} = b; \quad b_{12} = 0$$

i.e.,

$$\bar{a}_{11} = \bar{a}_{22} = \bar{a}; \quad \bar{a}_{12} = 0.$$

Such a situation actually occurs for a direction of propagation along the uniaxial axis in a crystal of uniaxial symmetry, since the sections of the index- and absorption-ellipsoids will be circular. Equations (16) and (15) then yield

$$\frac{D_y}{D_x} = \pm i; \quad \bar{v}^2 = \bar{a} \mp \bar{\Gamma}_z. \tag{19}$$

This means that the waves are right and left circularly polarised and if $\bar{v}_r$ and $\bar{v}_l$ be the complex velocities of the circularly polarised waves, then

$$\bar{v}_l^2 - \bar{v}_r^2 = 2\bar{\Gamma}_z. \tag{20}$$

The complex velocity $\bar{v}$ is related to the actual velocity v and the extinction coefficient κ by the usual relation

$$\bar{v} = \frac{c}{n - i\kappa} = v\left(1 + \frac{i\kappa v}{c}\right) \tag{21}$$

the terms containing the square of the extinction coefficient being negligible in magnitude. Introducing this in equation (20), we obtain to a high degree of approximation

$$n_r - n_l = \Gamma_z \cdot \frac{c}{v_m{}^3} = \rho_0 \tag{22}$$

$$\kappa_l - \kappa_r = \Gamma_z{}' \cdot \frac{c}{v_m{}^3} = \sigma_0 \tag{23}$$

where v_m is a mean velocity.

It will now be clear why Γ_z and $\Gamma_z{}'$ may be referred to as the parameters of optical rotation and circular dichroism corresponding to the direction of propagation z. The values of these parameters for a general direction of propagation s may be denoted by Γ_s and $\Gamma_s{}'$. The parameter of optical rotation has been shown to be a quadratic function of the direction cosines of propagation by virtue of (11) (*see* ref. 1, § 6); the same statement must therefore be true for the directional variation of the parameter of circular dichroism. If we lay off two radii vectores r_1 and r_2 parallel to the direction of propagation such that their lengths are given by

$$\frac{1}{r_1{}^2} = |\Gamma_s| = \rho_0 \cdot \frac{v_m{}^3}{c} \tag{24}$$

$$\frac{1}{r_2{}^2} = |\Gamma_s{}'| = \sigma_0 \frac{v_m{}^3}{c} \tag{25}$$

then (24) and (25) define respectively a surface of optical rotation and a surface of circular dichroism. Given these surfaces we may determine Γ_s and $\Gamma_s{}'$ for any direction of propagation, or alternatively the coefficients of circular birefringence and circular dichroism (ρ_0 and σ_0) for any direction.* (The sign to be attached may be supposed to be marked on the surface.)

§ 7. Analytical Presentation of the Method of Superposition

The states of polarisation of the waves propagated along any direction as well as their velocities and absorption coefficients may in principle be

* We take the mean velocity v_m for the direction in question to be given by equation (14) of reference 6.

obtained from equations (15) and (16). In practice, in order to be able to study the variation of these properties with direction the results would have to be cast into a simpler parametrical form—as Voigt[10] was compelled to do even in the case of inactive absorbing crystals. Now the propagation in an absorbing active crystal has been previously analysed without using the electromagnetic theory, by a method of superposition; this method, *when developed by the use of the Poincaré sphere*, leads to results which are automatically in such a parametrical form, and we shall show in the next section that these results can be modified so as to represent accurately the results of the electromagnetic theory. Unfortunately the modifications necessary can be perceived only if the results of the method of superposition, instead of being expressed in simple parametrical form, are expressed analytically by equations formally similar to those of the electromagnetic theory. For this purpose we shall in this section develop the consequences of the superposition method not by the geometrical methods of our previous paper,[2] but by analytical methods, along the lines followed by Jones[11,12] who has adopted a matrix calculus treatment.

A detailed physical description of the method of superposition may be found in references 6 and 2. Let $\vec{D}$ be the vibration at the plane z in the crystal.† The vibration which would obtain at the plane $(z + dz)$ if the crystal were transparent and inactive may be represented by $\vec{D} + \partial_1\vec{D}$. This vibration may be obtained by multiplying the components of $\vec{D}$ along the principal planes of linear birefringence Ox' and Oy' by the factors

$$\exp.\left(-i\frac{2\pi}{\lambda_0}n_1 dz\right)$$

and

$$\exp.\left(-i\frac{2\pi}{\lambda_0}\cdot n_2 dz\right) \text{ respectively.}$$

Then

$$\left.\begin{aligned} \partial_1 D_{x'} &= -i\frac{2\pi}{\lambda_0}n_1 D_{x'} \\ \partial_1 D_{y'} &= -i\frac{2\pi}{\lambda_0}n_2 D_{y'} \end{aligned}\right\} \tag{26}$$

† The displacement vector is here written as a two-dimensional vector since it lies on the wavefront. Also the vibration is taken as $\vec{D}e^{i\omega t}$, *i.e.*, in the present section the time factor is not included within the symbol for the displacement vector, which is therefore only a function of z.

This may be written symbolically as

$$\partial_1 \vec{D} = -i \frac{2\pi}{\lambda_0} dz\, [n]\, \vec{D}, \tag{27}$$

where $[n]$ represents a symmetric 2×2 matrix or tensor operator whose principal directions coincide with the principal planes of linear birefringence and whose principal values are n_1 and n_2—determined from an index-ellipsoid as for a transparent inactive crystal.

Similarly let $\partial_2 \vec{D}$ represent the partial increment to the initial vibration $\vec{D}$ when subjected to the infinitesimal operation of linear dichroism (corresponding to the thickness dz). Then we will have

$$\partial_2 \vec{D} = -\frac{2\pi}{\lambda_0} dz\, [\kappa]\, \vec{D} \tag{28}$$

where (κ) represents a symmetric tensor whose principal directions coincide with the principal planes of linear dichroism and whose principal values are κ_1 and κ_2 as given by an absorption ellipsoid [*see* eq. (12) of ref. 6].

Under the combined effects of linear birefringence and linear dichroism the partial increment $\partial' \vec{D}$ to the initial vibration $\vec{D}$ may be obtained by adding (27) and (28). We may then write

$$\partial' \vec{D} = -i \frac{2\pi}{\lambda_0} dz\, [\bar{n}]\, \vec{D} \tag{29}$$

the components of the symmetric tensor $[\bar{n}]$ being given by

$$\bar{n}_{ij} = n_{ij} - i\kappa_{ij} \tag{30}$$

where the n_{ij} and κ_{ij} are the components of $[n]$ and $[\kappa]$ respectively.

Thus the state of vibration $\vec{D} + \partial' \vec{D}$ at the plane $z + dz$ for an *inactive* absorbing crystal is given by writing equation (29) in full:

$$\left.\begin{aligned} \partial' D_x &= -i \frac{2\pi}{\lambda_0} dz\, (\bar{n}_{11} D_x + \bar{n}_{12} D_y) \\ \partial' D_y &= -i \frac{2\pi}{\lambda_0} dz\, (\bar{n}_{12} D_x + \bar{n}_{22} D_y) \end{aligned}\right\} \tag{31}$$

These equations correspond to eq. (29) of reference 6; the $\bar{n}_{ij}$ have the same significance as in that paper and are determined from the sections to the index- and absorption-ellipsoids.

Let $(\vec{D} + \partial_3\vec{D})$ represent the vibration obtained by subjecting the vibration $\vec{D}$ to the infinitesimal operation of rotation through an anticlockwise angle ρdz—where $\rho\ (\equiv \rho_0\pi/\lambda_0)$ is the optical rotatory power for the direction z as determined from a surface of optical rotation according to equation (24).

Then

$$\left.\begin{aligned} D_x + \partial_3 D_x &= D_x - \frac{\pi}{\lambda_0}\rho_0 dz . D_y \\ D_y + \partial_3 D_y &= \frac{\pi}{\lambda_0}\rho_0 dz D_x + D_y \end{aligned}\right\}. \tag{32}$$

It remains to consider the partial increment $\partial_4\vec{D}$ due to the operation of circular dichroism (corresponding to the passage dz): this operation consists in resolving the initial vibration $\vec{D}$ into its left and right circular components and multiplying the amplitudes of these components by $(1 - \pi/\lambda_0 \cdot \sigma_0 dz)$ and $(1 + \pi/\lambda_0 \cdot \sigma_0 dz)$; σ_0 is the coefficient of circular dichroism (corresponding to the difference in the extinction coefficients of the circularly polarised components) as obtained from a surface of circular dichroism according to equation (25). Clearly this operation differs from the operation of optical rotation (regarded as circular birefringence) *only in that the constant* $i\sigma_0$ *occurs in place of* ρ_0. Let $\partial''\vec{D}$ represent the partial increment to the initial vibration $\vec{D}$ due to the combined effects of optical rotation and circular dichroism (for the thickness dz). Then according to what has been said above $\partial''\vec{D}$ will be given by an expression which is similar to the expression for $\partial_3\vec{D}$ [which may be written down from equation (32)] except for the following difference: the constant ρ_0 will have to be replaced by $\bar{\rho}$ where

$$\bar{\rho} = \rho_0 + i\sigma_0. \tag{33}$$

In other words we will have

$$\left.\begin{aligned} \partial'' D_x &= -\frac{\pi}{\lambda_0} dz \cdot \bar{\rho} D_y \\ \partial'' D_y &= +\frac{\pi}{\lambda_0} \cdot dz \cdot \bar{\rho} D_x \end{aligned}\right\} \tag{34}$$

If $\vec{D} + d\vec{D}$ represents the vibration at the plane $z + dz$ in the optically active absorbing crystal then $d\vec{D}$ represents the total increment to the initial

vibration $\vec{D}$ under the combined effect of all the operations mentioned above. Thus $d\vec{D}$ is given by adding (31) and (34):

$$\left.\begin{aligned} dD_x &= -i\frac{2\pi}{\lambda_0}dz\,\{\bar{n}_{11}D_x + (\bar{n}_{12} - \tfrac{1}{2}i\bar{\rho})\,D_y\} \\ dD_y &= -i\frac{2\pi}{\lambda_0}dz\,\{(\bar{n}_{12} + \tfrac{1}{2}i\bar{\rho})\,D_x + n_{22}D_y\} \end{aligned}\right\}. \tag{35}$$

We wish to determine the *particular* states of polarisation of $\vec{D}$ for which the vibration is propagated unchanged with a specific complex velocity. From equation (14) we see that for such a disturbance we have

$$d\vec{D} = -i\frac{2\pi}{\lambda_0}dz\,.\,\bar{n}\vec{D} \tag{36}$$

Writing down the x and y components of (36) and comparing with (35) we finally obtain

$$\left.\begin{aligned} \bar{n} - \bar{n}_{11} &= (\bar{n}_{12} - \tfrac{1}{2}i\bar{\rho})\frac{D_y}{D_x} \\ \bar{n} - \bar{n}_{22} &= (\bar{n}_{12} + \tfrac{1}{2}i\bar{\rho})\frac{D_x}{D_y} \end{aligned}\right\} \tag{37}$$

which determine the characteristic states of polarisation D_y/D_x of the waves and their complex refractive indices $\bar{n}$—according to the method of superposition.

§ 8. SIMPLIFICATION OF THE RESULTS OF THE ELECTROMAGNETIC THEORY

The equations (37) deduced by the superposition method are formally similar to the equations (14) deduced by the electromagnetic theory. It may be shown that the quadratic equation for D_y/D_x which may be obtained from (37) will have coefficients proportional to the corresponding coefficients in equation (16)—the factor of proportionality being $(-c/2v_m{}^3)$. It follows that the states of polarisation of the waves as obtained by the electromagnetic theory are *identical* with those obtained by the superposition method. Hence the points representing these states of polarisation on the Poincaré sphere‡ may be determined exactly as described in ref. 2—the parameters (ϕ, ψ) which specify the states of polarisation being the same in both methods.

‡ The Poincaré sphere merely represents a method of mapping states of polarisation, and hence need not be used only in conjunction with methods of superposition.

In the second place, *because of the formal similarity between equations* (37) *and* (14), it is clear that for every equation derived by the method of superposition a corresponding equation obtainable from the electromagnetic theory may be written down by inspection—merely by replacing the symbols according to a scheme which transforms (37) to (14). In addition to the replacement scheme already used for this purpose in inactive absorbing crystals (ref. 6, § 8), we have only to add that the symbols ρ_0 and σ_0 have to be replaced $-2\Gamma_z$ and $-2\Gamma_z'$ respectively. Simple expressions for the refractive indices and absorption coefficients of the waves have been derived in ref. 2, § 9 *c*, by the superposition method (using the Poincaré sphere). The corresponding equations from the electromagnetic theory will be:

$$v^2 = \tfrac{1}{2}(a_1 + a_2) + \tfrac{1}{2}\sqrt{(a_1 - a_2)^2 + (2\Gamma_z)^2} \cdot \cos 2\phi. \tag{38}$$

$$\kappa \cdot \frac{2v^3}{c} = \tfrac{1}{2}(b_1 + b_2) + \tfrac{1}{2}\sqrt{(b_1 - b_2)^2 + (2\Gamma_z')^2} \cdot \cos 2\psi. \tag{39}$$

For directions of propagation not in the vicinity of an optic axis it can be shown (as in ref. 2, § 7) that the velocities and absorption coefficients of the waves may be determined from the index and absorption ellipsoids. It is only for directions in the vicinity of an optic axis that the complications caused by the ellipticity of the waves make their appearance; for such directions, since the birefringence is necessarily small, the difference between the above expressions and those given by the method of superposition will not in general be of any practical significance.

§ 9. The Singular Axes

Among the most remarkable of the properties of absorbing crystals—both inactive and active—is the possibility of the occurrence of so-called *singular axes*. A singular axis represents a direction for which the quadratic equation (16)—which determines the states of polarisation of the waves—has equal roots; the condition ($b^2 = 4ac$) under which this can obtain ensures at the same time the equality of the roots of the quadratic equation (15) which determines the velocities of the waves. Thus along a singular axis there is only *one* particular state of polarisation which can be propagated without change of form. In a previous paper[3] we have already discussed (with reference to the particular example of amethystine quartz) the following problems: the conditions under which it is possible for singular directions to exist in optically active crystals, the location of these axes, and the state of (elliptic) polarisation of the single wave that can be propagated along a singular axis. In the present paper, therefore, we confine ourselves to the following interesting query: what will happen when

a plane wave incident along a singular direction is in a state of polarisation *other* than that which *alone* can be propagated unchanged along that axis? In the particular case when the incident vibration is in a state of polarisation *orthogonal* to that of the wave which can be propagated unchanged along the singular axis, it had been supposed by Voigt[13] (in the case of inactive crystals) that the incident light would be totally reflected away, the reflection being partial in practical cases. That this is far from being the case has been shown experimentally by the author both in the case of iolite[7] (an optically inactive crystal) and amethyst[3] (an optically active crystal). The experimental results in the former case were readily explained[6] in detail (ref. 6, § 6 *b*) by a direct application of the method of superposition—according to which the incident vibration would obviously be propagated with a progressive change in its state of polarisation. *The same treatment and results apply* mutatis mutandis *to optically active absorbing crystals also.* Furthermore, it had been argued from considerations of continuity that the results given by the method of superposition could not really be contradictory to those of the electromagnetic theory. But how exactly the propagation of an arbitrary vibration along a singular axis could be directly handled by the electromagnetic theory was not apparent to the author till the publication of a recent paper by Jones[5]—in which is contained a new method of solving the electromagnetic equations in any anisotropic medium (transparent or absorbing). We present the method in a slightly modified form suitable for our present use.

§ 10. THE WAVE EQUATION FOR A PLANE WAVE PROPAGATED WITH CHANGE OF POLARISATION

In this section we wish to write down a suitable equation to represent a plane wave in a homogeneous medium, the wave being in general propagated with a progressive change in its state of polarisation. Firstly we shall take the vibration $\vec{D}$ at any plane z to be varying as $e^{i\bar{\omega}t}$ so that

$$\frac{\partial \vec{D}}{\partial t} = i\frac{2\pi}{\lambda_0} \cdot c\vec{D}. \tag{40}$$

Here the displacement vector has been written as a two-dimensional vector $\vec{D}$ since it is entirely transversal to the wave according to (3). As for the dependance of $\vec{D}$ on z, the form of the equation to be assumed is immediately suggested by equations (31) and (35) of § 7. The vibration at the plane $z + dz$ may be taken to be a linear *vector* function of the vibration

at the plane z (at the same instant), since it is *not* assumed that both are necessarily in the same state of polarisation. We may then write

$$\frac{\partial \vec{D}}{\partial z} = -i \frac{2\pi}{\lambda_0} [n'] \vec{D} \tag{41}$$

where $[n']$ is a two-by-two matrix operator. This resembles exactly the equation (36) which is satisfied by a usual plane wave of the form (4), but for the following difference: the refractive index $\bar{n}$ has been replaced by a 'refractive index tensor' $[n']$ in order that the same equation may represent a general plane wave propagated in a homogeneous anisotropic medium. This procedure is essentially in the same spirit as that of replacing the real refractive index by a complex quantity in order that the form of the usual equation to a plane wave in a transparent isotropic medium may still be retained to represent what is really a damped wave in an absorbing isotropic medium.

From (40) and (41), we obtain by differentiating with respect to t and z respectively

$$\frac{\partial^2 \vec{D}}{\partial t^2} = c^2 [n']^{-2} \frac{\partial^2 \vec{D}}{\partial z^2}. \tag{42}$$

The above equation represents the wave equation satisfied by the disturbance propagated with change of polarisation; it resembles the usual form of the wave equation to a plane disturbance except that in place of the square of the velocity we have the matrix operator $c^2 [n']^{-2}$.

As a simple example of a plane disturbance propagated with progressive change of polarisation we may remark that when a plane wave which is linearly polarised at a suitable azimuth is incident normally on a quarter-wave plate the incident linear vibration goes through progressive stages of elliptic polarisation and then emerges circularly polarised. This example also illustrates why, in usual cases, we need not directly seek general solutions representing disturbances propagated with change of polarisation: as will be evident in the example quoted, such a general solution is obtained by superposing the two characteristic plane wave solutions of the usual form (4)—*provided two such distinct solutions exist.*

§ 11. Second Method of Solving the Electromagnetic Equations

Our task is now to determine the refractive index tensor $[n']$ or alternatively the tensor $c^2 [n']^{-2}$. The result is sufficiently elegant to be stated

straightaway. Just as in an isotropic medium we have $n^2 = \epsilon$, where ϵ is the dielectric constant, in the present case it will be shown that

$$[n']^2 = [\epsilon], \tag{43}$$

where $[\epsilon]$ is the two-by-two matrix connecting $\vec{D}$ and $\vec{E}$:

$$\vec{D} = [\epsilon]\,\vec{E}. \tag{44}$$

Here $\vec{E}$ represents the 'projection' of the electric vector on the wavefront, being thus defined by the x and y components of the actual electric vector.

To establish (44) we note that the Maxwell's equations (2) may be written as

$$\frac{\partial^2 \vec{D}}{\partial t^2} = c^2 \frac{\partial^2 \vec{E}}{\partial z^2}. \tag{45}$$

The properties of the medium can be expressed in the form

$$c^2 \vec{E} = [A]\,\vec{D} \tag{46}$$

where [A] is a two-by-two matrix whose components may be written down from equation (13).

From (45) and (46) we have

$$\frac{\partial^2 \vec{D}}{\partial t^2} = [A]\,\frac{\partial^2 \vec{D}}{\partial z^2}. \tag{47}$$

Comparing with (42) we find that whatever be the state of polarisation of $\vec{D}$ a solution of the form (41) is possible with

$$c^2\,[n']^{-2} = [A]. \tag{48}$$

This relation is equivalent to (43) as may be seen by comparing (44) and (46). Thus the refractive index matrix $[n']$ can be determined§ for any direction of propagation from the relation

$$[n'] = c\,[A]^{-\frac{1}{2}} \tag{49}$$

We may briefly refer to the connection between the present method of solving the electromagnetic equations and that adopted in § 5. Though the wave equation (42) in general describes a disturbance propagated with

§ Only the physically significant square root of $[A]^{-1}$ is to be taken. For this and other mathematical questions which arise in the representation by matrix methods *see* Jones.[5,12]

a progressive change of polarisation this is not always the case. It will obviously reduce to the customary wave equation (for a homogeneously polarised disturbance) for those *particular* states $\vec{D}$ which satisfy the relation

$$[A]\,\vec{D} = \bar{v}^2\vec{D} \tag{50}$$

the complex velocity of the wave being $\bar{v}$. It will now be seen that the equations (13), previously used in § 5, are really the components of this vector equation, and as is shown by the procedure adopted there—the above equation is *usually* satisfied for two states of $\vec{D}$ (the eigenvectors of the matrix A) with *two* corresponding values of v^2 (the eigenvalues of the matrix A). For these particular states of $\vec{D}$ the equation (41) must also reduce to the form (36), *i.e.*, in the terminology of matrix calculus, these states of $\vec{D}$ are also the eigenvectors of the refractive index tensor $[n']$, the corresponding eigenvalues being the complex refractive indices of the waves.

From what has been said above, it is clear that a singular axis represents a special direction for which the matrix A has only *one* eigenvector, and correspondingly only *one* eigenvalue. This does not *in principle* lead to any difficulty in determining the refractive index matrix $[n']$ from (49), though of course the peculiar properties of [A] mentioned above are also carried over to $[n']$. On the other hand, when we adopt the method of superposition the refractive index tensor $[n']$ is considered as the sum of four parts each of which (it is assumed) may be directly determined in a simple fashion from the four corresponding surfaces defining the optical properties of the media. Thus whether we adopt the superposition method or the electromagnetic theory, no special difficulty arises in determining the refractive index tensor $[n']$ for propagation along a so-called singular direction. Along a singular direction—as, indeed, along any other direction—we can have plane disturbances which are propagated with a progressive change of polarisation: the speciality about a singular direction is, however, that such a disturbance cannot in turn be regarded as the sum of two plane waves of constant polarisation—there being only *one* wave of the latter type.

The author owes a debt of gratitude to Prof. Sir C. V. Raman, at whose instance the initial experimental investigations on amethyst and iolite were undertaken by the author, and without whose encouragement the subsequent theoretical investigations on the properties of absorbing crystals would not have been possible.

§ 12. SUMMARY

The propagation of light along an arbitrary direction in an absorbing active crystal is solved by extending the treatment previously given[1] for transparent active crystals—the index tensor and the modified gyration tensor being replaced by corresponding tensors with complex components. The two waves are in general elliptically polarised, in states identical with those given by a method of superposition[2]; their velocities and absorption coefficients are likewise simple functions of the parameters which specify these states of polarisation on the Poincaré sphere.

Attention is drawn to the propagation along any singular direction (in active or inactive crystals) where only *one* homogeneously polarised plane wave solution is obtained—and not two. A more general theoretical approach[5] becomes necessary to establish—in agreement with experiment—that other solutions also exist, representing plane disturbances propagated with a progressive change of polarisation. For such a disturbance the wave equation satisfied by the displacement vector differs from the usual form only in that the square of the velocity has to be regarded as a tensor operator.

§ 13. REFERENCES

1. Pancharatnam .. *Proc. Ind. Acad. Sci.*, 1956, **43 A**, 247.
2. ——— .. *Ibid.*, 1957, **46 A**, 280.
3. ——— .. *Ibid.*, 1958, **47 A**, 201, 210.
4. Forsterling .. *Gottinger Nachr.*, 1912, p. 207.
5. Jones .. *J. Opt. Soc. Am.*, 1956, **46**, 126.
6. Pancharatnam .. *Proc. Ind. Acad. Sci.*, 1955, **42 A**, 86.
7. ——— .. *Ibid.*, 1955, **42 A**, 235.
8. Page and Adams .. *Electrodynamics*, Von Nostrand, 1949, Ch. 8.
9. Born .. *Optik*, Springer, 1933, p. 414.
10. Pockels .. *Lehrbuch der Kristalloptik*, Teubner, 1906, pp. 386–99.
11. Jones .. *J. Opt. Soc. Am.*, 1942, **32**, 486.
12. ——— .. *Ibid.*, 1948, **38**, 671.
13. Voigt .. *Ann. Physik.*, 1908, **27**, 1002.

APPENDIX

(*a*) *Derivation of the constitutive relation in the form* (6).—The relation between **D** and **E** for an optically active absorbing crystal may be written as

$$\mathbf{D} = (\epsilon)\,\mathbf{E} + \mathbf{P}_1 + \mathbf{P}_2. \tag{46}$$

The portion of the induced polarisation $\mathbf{P}_1$ contributes to the mean absorption and the linear dichroism, while $\mathbf{P}_2$ leads to optical rotatory power and circular dichroism. These are given by the expressions (*see* ref. 9):

$$\mathbf{P}_1 = -\,i\,(\epsilon')\,\mathbf{E};\quad \mathbf{P}_2 = i\,\overline{\mathbf{G}} \times \mathbf{E}. \tag{47}$$

For usual values of absorption and optical activity $\mathbf{P}_1$ and $\mathbf{P}_2$ are so small that terms of the second order in $\mathbf{P}_1$ and $\mathbf{P}_2$ may be ignored. Hence to this degree of approximation $\mathbf{P}_1$ and $\mathbf{P}_2$ may be expressed as functions of **D** by substituting in (47) an approximate value of **E** obtained from (46), *viz.*, $\mathbf{E} = (\epsilon^{-1})\,\mathbf{D}$. We then obtain

$$\begin{aligned}(\epsilon)\,\mathbf{E} &= \mathbf{D} - \mathbf{P}_1 - \mathbf{P}_2\\ &= \mathbf{D} + i\,(\epsilon'\epsilon^{-1})\,\mathbf{D} - i\,(\mathrm{R}\epsilon^{-1})\,\mathbf{D}\end{aligned}$$

where the operator $\overline{\mathbf{G}}\times$ is replaced by an antisymmetric matrix operator R (*see*, *e.g.*, ref. 8, eq. 85·2). We then obtain

$$\mathbf{E} = (\epsilon^{-1})\,\mathbf{D} + i\,(\epsilon^{-1}\epsilon'\epsilon^{-1})\,\mathbf{D} - i\,(\epsilon^{-1}\mathrm{R}\epsilon^{-1})\,\mathbf{D}. \tag{48}$$

If we choose axes of co-ordinates X, Y, Z along the principal axes of the dielectric tensor (ϵ) it can be shown by actual matrix multiplication that the operator $(\epsilon^{-1}\mathrm{R}\epsilon^{-1})$ is equal to the operator $1/c^2 \cdot \overline{\Gamma}\times$, where

$$\overline{\Gamma}_{\mathrm{X}} = \frac{c^2}{\epsilon_{\mathrm{Y}}\epsilon_{\mathrm{Z}}} \cdot \overline{\mathrm{G}}_{\mathrm{X}};\quad \overline{\Gamma}_{\mathrm{Y}} = \frac{c^2}{\epsilon_{\mathrm{Z}}\epsilon_{\mathrm{X}}} \cdot \overline{\mathrm{G}}_{\mathrm{Y}};\quad \overline{\Gamma}_{\mathrm{Z}} = \frac{c^2}{\epsilon_{\mathrm{X}}\epsilon_{\mathrm{Y}}} \cdot \overline{\mathrm{G}}_{\mathrm{Z}}. \tag{49}$$

Comparing (48) with (6) and (8) we see that the vector of optical activity $\overline{\Gamma}$ is given by the above equations, while the index and absorption tensors are given by

$$(a) = c^2\,(\epsilon^{-1});\quad (b) = c^2\,(\epsilon^{-1}\epsilon'\epsilon^{-1}). \tag{50}$$

It must be remembered that $\overline{\mathbf{G}} = (\bar{g})\,\mathbf{s}$, where $(\bar{g})$ is the complex gyration tensor and **s** the wave-normal. If we write this relation in full (*see* ref. 9) it follows from (49) that the relation between the components of the tensor

2

of optical activity ($\bar{\gamma}$) and the components of the complex gyration tensor ($\bar{g}$) will be given by

$$\bar{\gamma}_{1m} = \frac{c^2}{\epsilon_Y \epsilon_Z} \cdot \bar{g}_{1m}; \quad \bar{\gamma}_{2m} = \frac{c^2}{\epsilon_Z \epsilon_X} \bar{g}_{2m}; \quad \bar{\gamma}_{3m} = \frac{c^2}{\epsilon_X \epsilon_Y} \cdot \bar{g}_{3m} \tag{51}$$

$$(m = 1, 2, 3).$$

(*b*) *An error in Jones' paper.*—In a paper by Jones[5] dealing with propagation in anisotropic media, the constitutive relation between **D** and **E** for an optically active medium has been expressed in a form which has been thought to be equivalent to that given above [eqs. (46) and (47)] but is really quite different from it. In particular the term $\mathbf{P}_2$ in (47) has been assumed by Jones to be given by

$$\mathbf{P}_2 = \frac{c}{\omega} \{(\bar{g}) \nabla\} \times \mathbf{E} \tag{52}$$

where ∇ is the gradient operator and ($\bar{g}$) the complex gyration tensor. That this relation is quite different from the customary form (47) may be easily seen in the particular case when we seek plane wave solution of the usual form (4) so that $\nabla = i\omega\, \bar{n}/c \cdot \mathbf{s}$. The relation (52) then becomes

$$\mathbf{P}_2 = i\bar{n}\,(\mathbf{G} \times \mathbf{E}). \tag{53}$$

This is of course vitally different from the expression given in (47) since according to (53) the refractive index $\bar{n}$ which is one of the unknowns to be determined will itself be involved in the constitutive relations; more generally the refractive index tensor (n') will itself be involved in the matrix [A] occurring in (43). Accordingly, the relations given by Jones expressing his 'N-matrix' in terms of the complex dielectric and gyration-tensors are in error.

Reprinted from "The Proceedings of the Indian Academy of Sciences", Vol. XLVIII, No. 6, Sec. A, 1958

PERCUSSION FIGURES IN CRYSTALS

BY SIR C. V. RAMAN

(*Memoir No.* 111 *from the Raman Research Institute, Bangalore*-6)

Received November 29, 1958

1. INTRODUCTION

WHEN a polished sphere of hard steel is dropped from a height on the plane smooth surface of a thick slab of glass and rebounds therefrom, the impact results in a beautiful effect which is seen within the glass around the region of contact between the sphere and the slab. The phenomenon was described and illustrated by a set of photographs in a communication published many years ago under the title " Percussion Figures in Isotropic Solids " in *Nature* of the 9th October 1919. When an optical test-flat is laid on the surface of the glass slab over the region of impact, the nature of the permanent deformations resulting from the impact is very clearly revealed by the interference patterns seen in monochromatic light between the two surfaces. Photographs of these patterns were published in an article by the author on " The Optical Study of Percussion Figures " which appeared in the *Journal of the Optical Society of America* for April 1926. A more detailed study of the phenomena including especially a quantitative comparison between the facts of observation and the consequences of Hertz's well-known theory of impact appeared shortly afterwards in a paper by one of the author's collaborators in the *Indian Journal of Physics.*

It had long been the author's intention to investigate by similar methods the results of the impact of a steel sphere on single crystals of various materials. Only recently however did it become possible for this programme to be taken up and the present communication reports the results. The materials investigated were quartz, calcite, barytes, and felspar, and to make the study more complete, a few polycrystalline solids found in nature were also investigated. Photographs of the percussion figures obtained in the various cases are reproduced with the paper. The cases studied most fully are those of quartz and calcite. The results obtained with quartz are illustrated by Figs. 1 to 6 in Plate XXV and Figs. 1 to 6 in Plate XXVI. Plates XXVII, XXVIII and XXIX illustrate the results obtained with the other solids.

2. PERCUSSION FIGURES IN GLASS

Three photographs reproduced respectively as Figs. 1, 2 and 3 in Plate XXVII exhibit the results of the impact of a steel sphere on the surface

of a thick glass plate. Simple inspection reveals that the external surface of the slab remains uninjured by the impact except along a ring or annulus with a sharply defined inner boundary. Within the annulus is a clear circular area which evidently represents the region where the sphere and plate come into contact during the impact. From the ring or annulus a fracture spreads obliquely inwards within the glass in the form of a surface of revolution. Fig. 1 in the Plate is a lateral view of the fracture seen through the edge of the plate; its mirror-image as seen by internal reflection at the surface of the plate also appears in the photograph. Interference rings are exhibited by the fracture both by reflected and by transmitted light; their configuration indicates that the separation between the two faces of the fracture is a maximum at the surface of the plate and diminishes progressively to zero at its termination in the interior (Fig. 2). The interference pattern which is observed when a test-plate is laid on the surface is reproduced as Fig. 3. It shows the area of the percussion figure to consist of three distinct regions: (*a*) a circular central area which remains plane and undeformed, but exhibits a small permanent lowering of its level; (*b*) an annular region of fracture showing severe injury and (*c*) a sudden elevation of the surrounding area which slopes gradually down to the general level of the surface at the outer margin of the percussion figure; the last-mentioned feature is clearly a consequence of the internal fracture.

While the above is a general description, there are also other subsidiary features deserving of mention. When the circumstances of the impact are such that the pressure exerted by the sphere on the plate just suffices to produce a fracture, the circular ring-crack is sharp and well defined. On the other hand, when the velocity of the impinging sphere in relation to its size exceeds the minimum needed to cause a fracture, we have instead of a single circular crack, several cracks forming an annulus, the external diameter of which corresponds to the maximum area of contact, while the inner diameter is noticeably smaller. Subsidiary cracks also spread out in various directions along the surface of fracture interrupting its smooth continuity. Several of these can be seen in Fig. 2 in Plate XXVII.

3. PERCUSSION FIGURES IN POLYCRYSTALLINE SOLIDS

Several naturally occurring solids are known which are polycrystalline in texture and which by reason of their mechanical strength find useful applications as building materials. Marble is one of the best known of them. It is however rather coarsely crystalline in its texture. Of greater interest from our present point of view is a material which may be described as a fine-grained limestone exhibiting varied colours and of which enormous quanti-

ties are quarried in Southern India for use as flooring tiles. According to the locality of its origin, the material is variously known as Cuddappah stone, Shahabad stone, Tandur stone, etc. It is easily split into slabs of the desired thickness and the surface of the slabs can be worked to any desired degree of perfection. Indeed, it can be polished so perfectly as to give regular interference fringes when a test-flat is laid on it.

Figs. 4, 5 and 6 in Plate XXVII show the highly characteristic features of the percussion figures shown by the Tandur stones. Simple inspection by reflected light reveals a spherical dimple covering the area of contact between the sphere and the slab. A closer examination reveals a concentric set of sharply defined circular cracks lying on the slopes of the depression. One also observes a set of radial cracks spreading outwards from the edge of the cup-shaped depression. These features can be recognized in Fig. 4 of the Plate. The interference patterns observed with a test-plate are of a totally different nature within the area of the depression and in the region outside it. They form a set of closely spaced circular rings within the depression (Fig. 6), while in the outer area they form a roughly hexagonal pattern the edges of which evidently correspond to the radial cracks spreading outwards (Fig. 5 in the Plate).

A polished surface of marble shows very clearly the shallow depression produced by the impact of a steel sphere. Simultaneously also, the impact results in a visible breaking up of the texture of the marble, so much so that the entire area affected by it acquires a frosted appearance. We do not, however, observe either the circular or the radial cracks which are so conspicuous a feature in the case of the Tandur stones. The depressions in marble can also be made visible by laying a test-plate on the surface. This is shown in Figs. 3 and 4 in Plate XXVIII which refer to two different samples of marble.

Broadly speaking, it may be said that coarsely grained solids exhibit effects in impact which are similar to those exhibited by marble. This is illustrated by Figs. 5 and 6 in Plate XXIX which represent the deformations produced by impact as revealed by a test-plate. Fig. 5 refers to the case of a black limestone akin to marble, while Fig. 6 represents the percussion figure in dolerite which is a hard and tough building stone capable of taking a high polish.

4. Percussion Figures in Quartz

Quartz does not exhibit any regular cleavages and shows only irregular fractures when broken under a hammer. In this respect it resembles glass.

Prima facie therefore, we may expect quartz to exhibit percussion figures broadly resembling those observed with glass but differing therefrom in detail by reason of the fact that quartz is a crystal with anisotropic physical properties. These expectations are borne out by the facts. There are however some additional features of a rather surprising and unexpected nature.

In studying the percussion figures of crystals, it is obviously necessary that the surface on which the impact occurs is optically plane and has a specified orientation in relation to the axes of crystal symmetry. It is also necessary that the specimen employed is sufficiently massive in relation to the size of the impinging sphere. These conditions are fortunately attainable without much difficulty in the case of quartz. Specimens of adequate size and of reasonably good quality are obtainable which admit of being cut and polished in any desired direction. Three specimens of quartz were employed in the present investigation. One of them was a massive crystal weighing several kilograms. On this, a large surface about 20 cm. × 10 cm. normal to the optic axis was cut and polished. A second face of nearly the same size parallel to a natural prismatic face was also ground and polished. Though the material was not of the highest optical quality, there were regions abutting on the two prepared surfaces which were quite clear and transparent and of sufficiently great extension in relation to the size of the percussion figures. The results observed are therefore worthy of confidence. The second specimen was a slab of crystal quartz about 20 cm. × 15 cm. and 5 cm. thick, the faces of which were normal to the optic axis. One of the faces was carefully ground and polished. The third specimen employed was a crystal in its natural form exhibiting six prismatic faces, besides three large and three small rhombohedral faces. All these were ground and polished to perfection. The optical quality of the material was however rather poor.

Most of the detailed studies made refer to the case in which the surface on which the impact occurs is a basal section. Indeed, all the photographs reproduced in Plates XXV and XXVI refer to this situation. The reason for the choice made is obvious; for, since the impact in this case is along the optic axis of the crystal, we should expect the percussion figure to exhibit the maximum degree of symmetry. This expectation is borne out by the facts. Some visual observations were however also made for percussion figures on the prismatic and rhombohedral faces.

It is instructive to compare the photographs reproduced as Figs. 1 to 6 in Plate XXV and as Figs. 1 to 6 in Plate XXVI which, as already stated, represent the percussion figures of quartz on a surface normal to the optic

axis with Figs. 1 to 3 in Plate XXVII which show the percussion figures in glass. There are certain general similarities, but the differences are also striking and noteworthy. Alike in glass and in quartz, the impact results in the formation of cracks which commence on the surface around the margin of the area of contact and spread inwards into the material. But here all resemblance ends. The surface cracks in quartz are exceedingly fine and are seen under a magnifier as hair-lines clearly separated from each other, and there is otherwise no visible damage to the surface. The assembly of surface-cracks is not circular in shape but resembles a hexagon with rounded edges. The fracture also spreads into the interior of quartz much more steeply than in the case of glass. This will be evident on a comparison of Fig. 4 in Plate XXVI which is an attempt to photograph the percussion figure in quartz as seen laterally with Fig. 1 in Plate XXVII which is the corresponding figure for glass. The fracture-surface in the interior of quartz rarely shows any interference colours either by reflection or in transmission, from which it is clear that the two faces of fracture are very close to each other. By reason of its steepness, however, the fracture can cut off the transmitted light and then appears dark. This is clearly seen in Figs. 1, 2 and 3 of Plate XXV and also Fig. 3 of Plate XXVI. As light tilt of the fracture-surface in relation to the direction of the incident light however makes a large change in this respect (see Figs. 2 and 3 in Plate XXV). The fracture-surface can also be seen and photographed by reflected light (Figs. 1 and 2 in Plate XXVI).

As will be evident from the photographs reproduced in the two Plates, the percussion figure inside quartz exhibits only trigonal symmetry, and not hexagonal symmetry. The photographs, however, do not adequately picture the actual configuration of the fracture. The shape of the latter is perhaps best described as resembling a three-cornered hat. In other words, the fracture takes the form of three ribs with three flaps separating them. The flaps are much more easily seen by transmitted and reflected light than the ribs. Another and most interesting feature is the fine structure exhibited by the fracture within the quartz. One observes a radial fibrous structure and crossing this we have also a circumferential ribbing consisting of a great number of concentric circles which run all the way down from the surface to the termination of the fracture. The circumferential ribbing is most clearly seen in Fig. 1 of Plate XXV and not quite so clearly in Figs. 1 and 2 of Plate XXVI. The radial structure is very clearly manifest in the two latter photographs.

As the fracture within the quartz exhibits only trigonal symmetry, the elevation of the surrounding surface may naturally also be expected to exhibit

the same symmetry. This is manifest from the interference patterns observed when a test-plate is laid on the surface (see Figs. 4, 5 and 6 in Plate XXV and Figs. 5 and 6 in Plate XXVI). The trigonal symmetry is particularly clear in the second of the two patterns reproduced in Fig. 6 in Plate XXV.

It would not be profitable to attempt here any explanation or interpretation of the various features set out above. That the percussion figure for impact on a basal plane, in other words along the optic axis of quartz, exhibits only trigonal symmetry and not hexagonal symmetry is interesting, but not altogether surprising. For, we know that the optic axis of quartz is a three-fold axis and not a six-fold axis of symmetry of the structure.

With regard to the observations which have been made of the percussion figures on the prismatic and rhombohedral faces, it will suffice here to remark that they do not exhibit the trigonal symmetry described and illustrated in Plates XXV and XXVI. In other respects, however, they exhibit features very similar to those set forth above. Each of the two cases exhibits special features of its own. But the observations made are not sufficiently numerous and trustworthy to justify a detailed description or discussion of the same.

5. PERCUSSION FIGURES IN CALCITE

A calcite crystal in its natural form as a rhomb with edge-lengths 8 cm. × 6 cm. × 4 cm. was employed in the study. Though the material was not of the highest optical quality throughout its volume, a region adjacent to one of the larger faces (8 cm. × 6 cm.) was clear and transparent to a considerable depth. Accordingly, this particular face was ground and polished and utilized for the production of the percussion figures. One of the adjacent faces (6 cm. × 4 cm.) was also ground and polished so as to permit of the interior of the crystal being viewed through it. In view of the small size of the specimen, the steel sphere employed was also small, about a centimetre in diameter. It had a highly polished surface and was dropped on the polished face of the rhomb from different heights at eight different points sufficiently remote from each other to ensure the absence of any mutual interference.

Percussion figures are observed around each of the chosen points of impact. They are conspicuous when the faces of the rhomb are viewed by reflected light, and they can also be seen to penetrate into the crystal when viewed longitudinally and also when observed laterally through the adjacent polished face. All the eight figures exhibit a general similarity with each other, though their sizes are different, since they diminish with the velocity of impact. They are all similarly orientated on the face of the rhomb and

the configuration of each is symmetric about a plane normal to the face of the crystal which makes equal angles with its two rhombohedral edges. But they do not exhibit any symmetry with respect to a perpendicular plane and are indeed strikingly different on the two sides of it. The actual region of contact between the sphere and the crystal can be readily recognized by the fact that it appears relatively dark as seen by reflected light in comparison with the brilliant display of colours all round. It is also evident on a careful examination that it is slightly depressed in level below the original surface of the crystal.

Viewed by reflected daylight, the percussion figures exhibit a complex pattern of interference colours and fringes. The interferences are also conspicuously visible in monochromatic light. A photograph of the same is reproduced as Fig. 1 in Plate XXVIII. When an optical test-flat is laid on the calcite and allowed to settle down so that it comes into close contact with the crystal face, the interference pattern seen differs very little indeed from that observed without the test-plate. This fact is evident on a comparison of Fig. 1 in Plate XXVIII with Fig. 2 in the same Plate, the latter representing the fringes seen when the test-plate has been laid on. If there are any fringes between the test-plate and the crystal, they are so faint and broad that their superposition on the closely spaced fringes in the interior makes little difference to what is actually observed.

It is well known that calcite has three planes of easy cleavage parallel to the three faces of the rhombohedron and that it also possesses three glide planes each containing one of the rhombohedral edges and equally inclined to the other two. An examination of the percussion figures in calcite reveals that the two cleavage planes intersecting the face on which the impact occurs play an important role in determining the results of the impact. It is observed that on either side of the area of contact between the sphere and the crystal, two cleavages making an acute angle with each other develop and extend outwards from the edges of that area. These cleavages are clearly visible on the face of the crystal, and they sharply limit the areas within which the fracture develops and spreads inwards into the crystal. Indeed, the pattern of reflected colours as seen in daylight and the interference pattern as seen by monochromatic light manifest themselves wholly within the limits set by these two cleavages. These features can be clearly recognized in Fig. 1 of Plate XXVIII. A further and most interesting feature is the appearance of a whole series of parallel lines outside the region of contact and only one side of it and which extend over a considerable area of the pattern. These lines are equally inclined to the two sets of cleavages and

cut across the curved interference bands exhibited by the fracture. We may explain them as due to glides occurring in the crystal along the direction of the third rhombohedral edge. The direction of the glide deviates greatly from the normal to the crystal face on which the impact occurs and hence also from the direction of impact. This may explain why the fracture-surface within the crystal extends much more in the forward direction than in the backward.

6. Percussion Figures in Barytes

The specimen of barytes employed was translucent with a faint bluish tinge. Its original shape as a crystal was that of a parallelopiped, four of whose faces were nearly rectangular, while the other two deviated sensibly from that shape. An artificial face of larger dimensions was obtained by cutting through the specimen in a direction making 45° with the rectangular faces; it was then ground and finely polished. In view of the small size of the specimen, the steel ball used had also to be quite small, only about 6 mm. in diameter. It was dropped from various heights and the percussion figures thus obtained were critically examined.

The effects produced by the impact may be summed up as follows. Firstly, we have a visible dimple or depression left on the surface. There is also a general disturbance of the level of the crystal face around the point of impact, and this is not spherically symmetrical around the region of impact. We have also noteworthy changes in the interior of the crystal which result in a brilliant reflection appearing in certain directions, while it appears dark in other directions. This effect varies greatly with the direction of illumination and the direction of observation, some areas appearing dark in certain circumstances and others bright and *vice versa.* A remarkable lack of symmetry is exhibited by the effect. Merely turning round the specimen through 180° under oblique illumination and normal observation results in all the areas which appear dark becoming bright and *vice versa.*

The internal reflections exhibited by one of the patterns is reproduced as Fig. 5 in Plate XXVIII. As both the illumination and the observation were normal, the special features mentioned above are not exhibited. Seen under oblique illumination, the figure resembles a butterfly with outstretched wings and tail, the head being bright and the tail and wings dark or *vice versa.* The changes in the interior—presumably in the nature of fractures—are accompanied by changes in the exertnal level of the surface. These latter are revealed by the interference fringes observed when a test-plate is put on the specimen. This is illustrated by Fig. 6 in Plate XXVIII.

This has the same orientation as Fig. 5, the latter being the figure as seen by reflected light.

7. The Percussion Figures in Felspar

Specimens of felspar of large size are readily available exhibiting the natural faces of the crystal. These faces can be ground and polished so as to be optically plane. Thus, felspar lends itself readily to an exact study of the results of impact on crystal surfaces of known orientation. Unfortunately, however, it is not easy to obtain crystals of any appreciable size which are *optically clear*. The observations have accordingly to be confined to those features resulting from impact which are *externally* observable. This makes it rather difficult to make any detailed analysis of the results of impact in terms of crystal structure and orientation.

A highly characteristic effect exhibited by felspar and produced by the impact of a steel sphere on a natural crystal face is illustrated in Fig. 1 in Plate XXIX. Viewed by oblique illumination, the area of contact appears as a dark crescent surrounded by a bright crescent beyond it. Careful examination shows that the bright crescent is an internal reflection. Merely turning round the specimen through 180° causes the reflection to disappear. The reflection presumably arises from a fracture in the interior of the felspar produced by the impact. This naturally also results in an elevation of the surface and this is revealed by placing a test-plate on the surface. We then observe an asymmetric distribution of interference rings around the region of the impact. Five such patterns are seen in Figs. 2 and 3 in Plate XXIX. It will be noticed that the interference patterns are all similarly orientated, indicating that the phenomenon is definitely related to the crystal structure of the material. Fig. 4 is a photograph of the interference pattern exhibited by the percussion figure on a different natural face of a felspar crystal. It will be seen that it has an altogether different character from those reproduced in Figs. 2 and 3 in the same Plate.

* * * * *

The preparation of the material employed in these studies and of the pictures which illustrate the paper required much painstaking labour. The work was undertaken and successfully carried out by my research assistant Mr. J. Padmanabhan.

* * * * *

SUMMARY

The impact of a hard steel sphere on the optically polished surface of a solid results in a permanent deformation of the surface and also produces fractures or rearrangements in the interior of the solid. These have been studied with crystals of quartz, calcite, barytes and felspar and also with a few polycrystalline solids. The results of the study show clearly that the nature of the percussion figure exhibited by a solid is a characteristic property of the material and is related to its inner structure and symmetry. The paper is illustrated by a series of photographs.

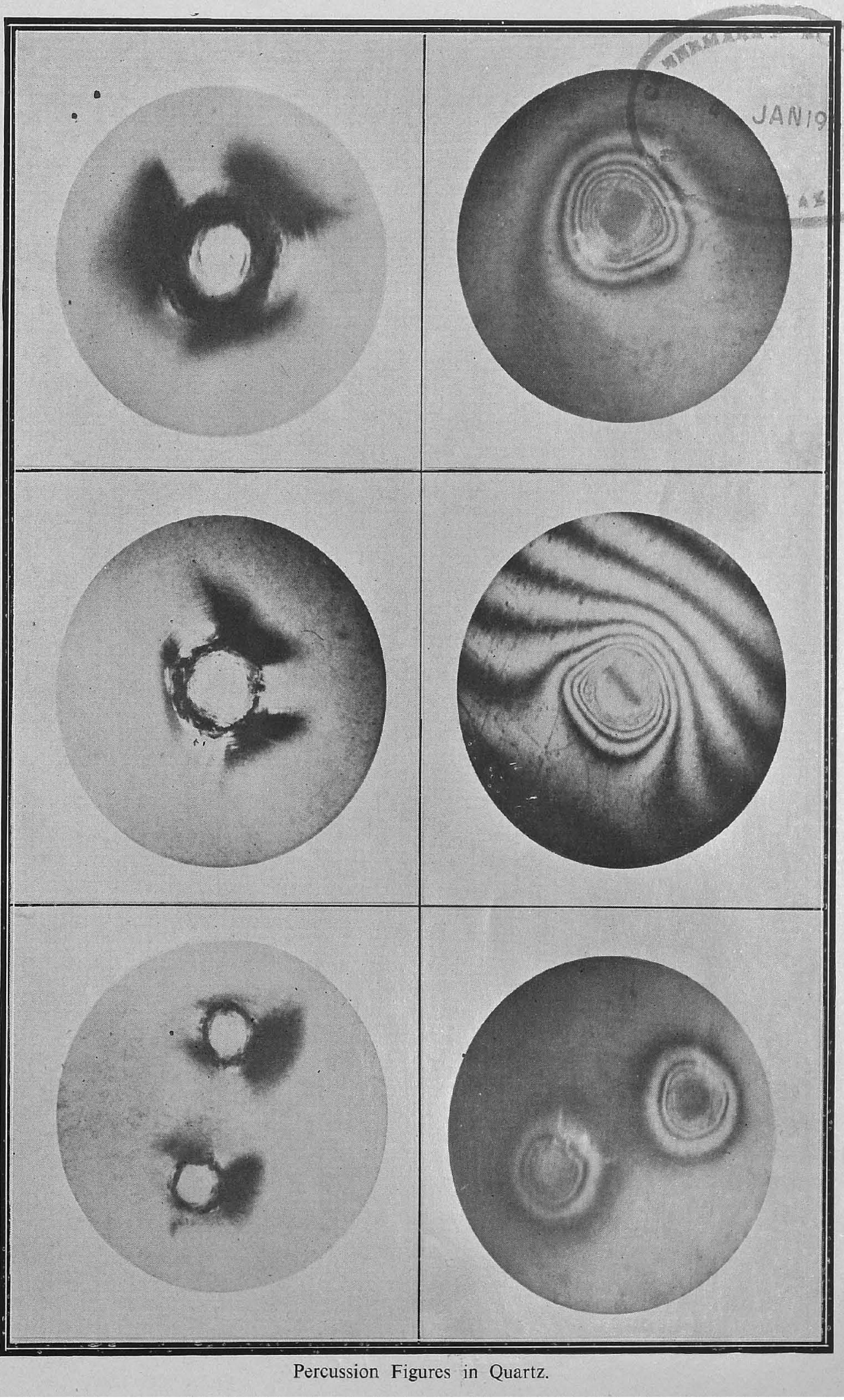


FIG. 1 FIG. 4

FIG. 2 FIG. 5

FIG. 3 FIG. 6

Percussion Figures in Quartz.

FIG. 1

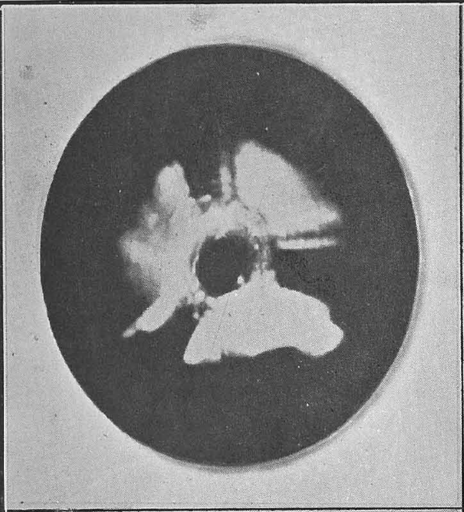


FIG. 2

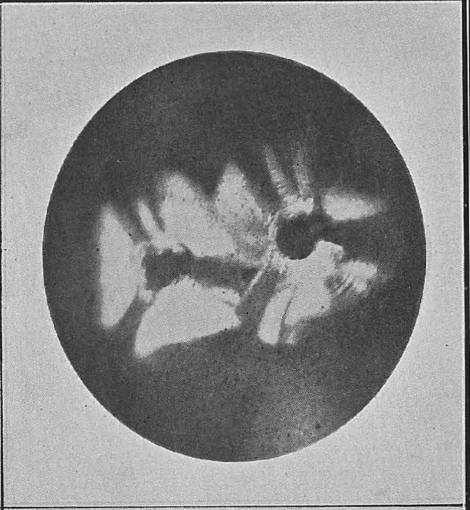


FIG. 3

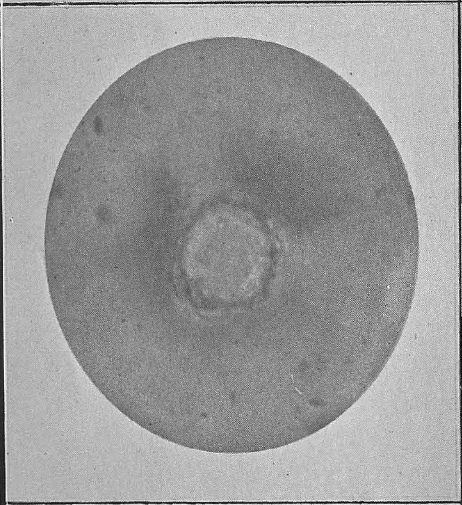


FIG. 4

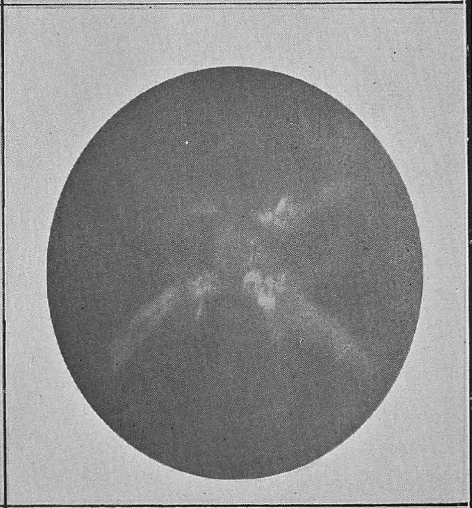


FIG. 5

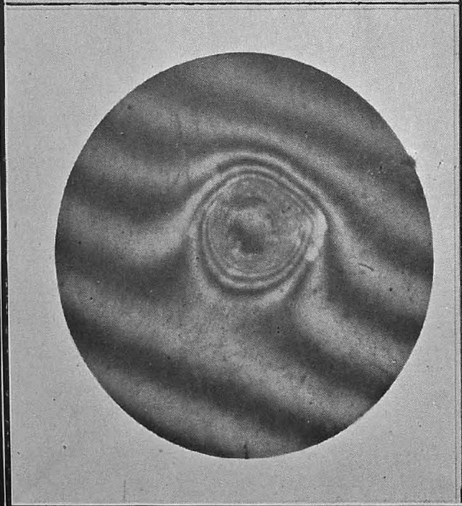
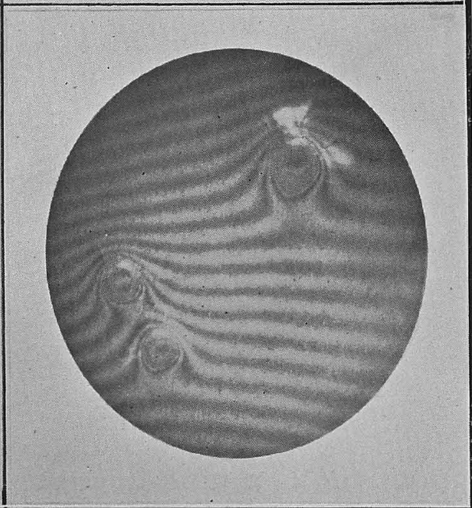


FIG. 6



Percussion Figures in Quartz.

CONNEMARA PUBLIC LIBRARY
4 JAN 1965

FIG. 1

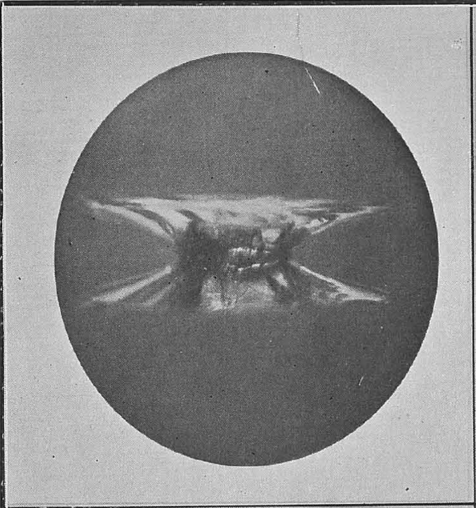
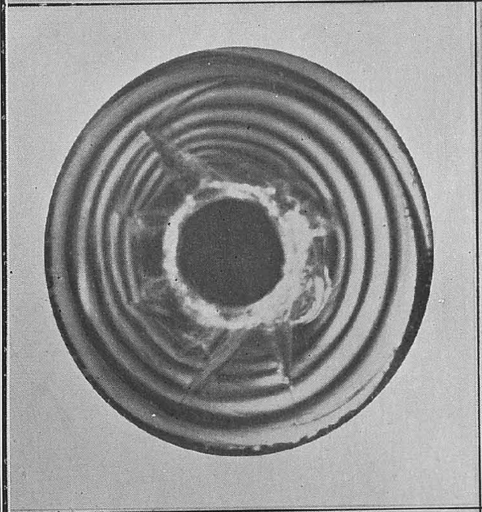


FIG. 4

FIG. 2



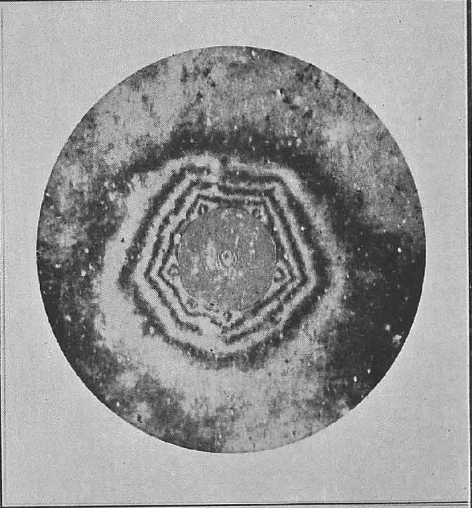


FIG. 5

FIG. 3

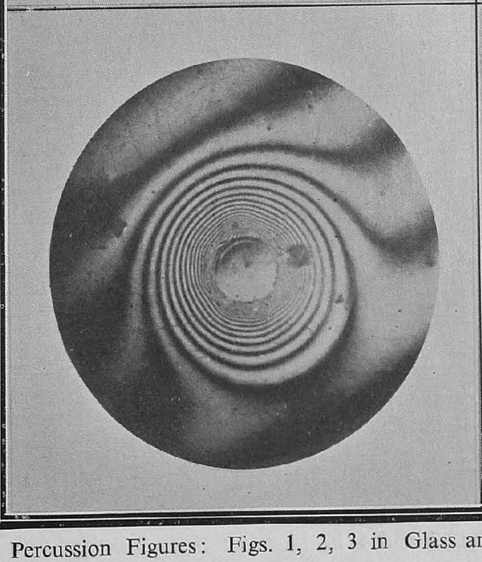


FIG. 6

Percussion Figures: Figs. 1, 2, 3 in Glass and Figs. 4, 5, 6 in Polycrystalline Limestone.

FIG. 1

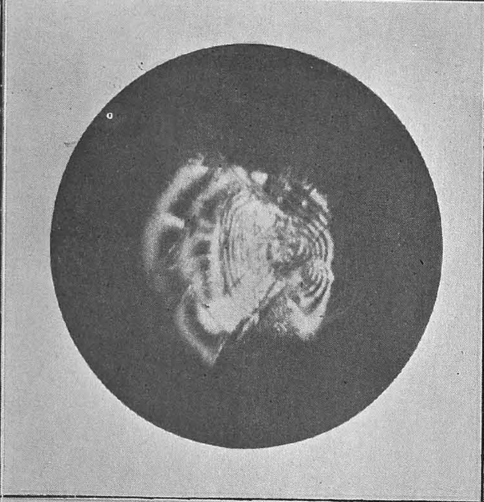


FIG. 2

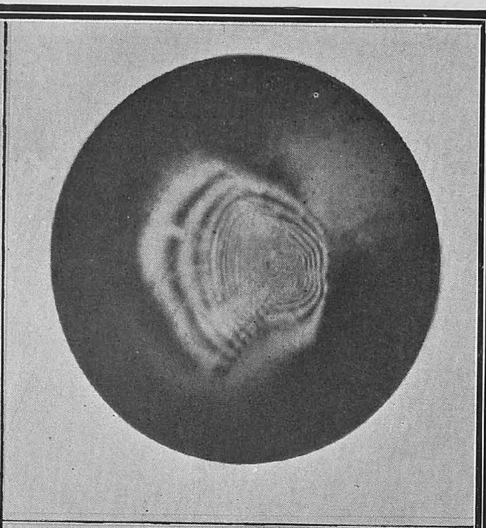


FIG. 3

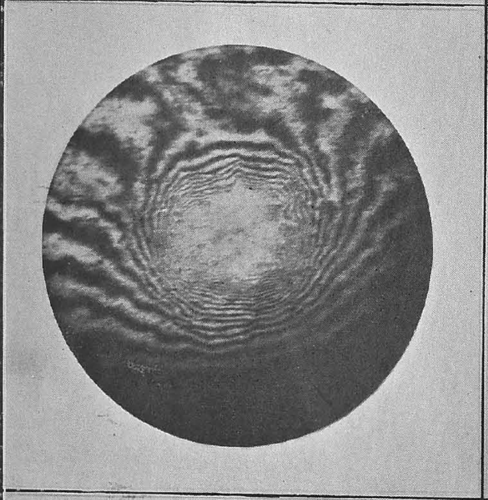


FIG. 4

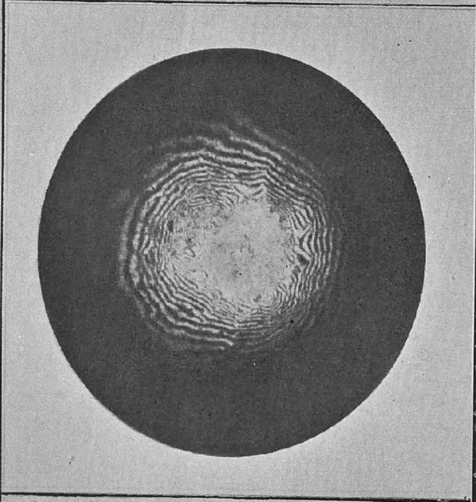


FIG. 5

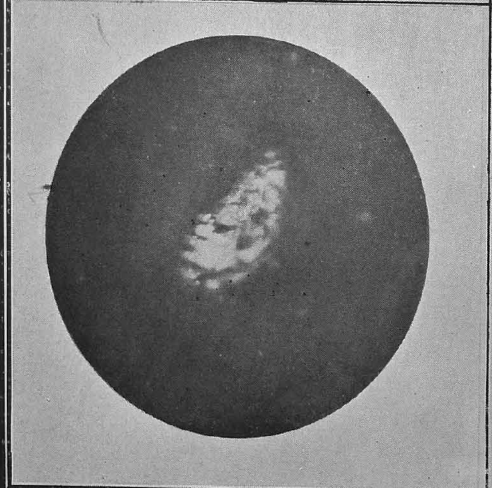
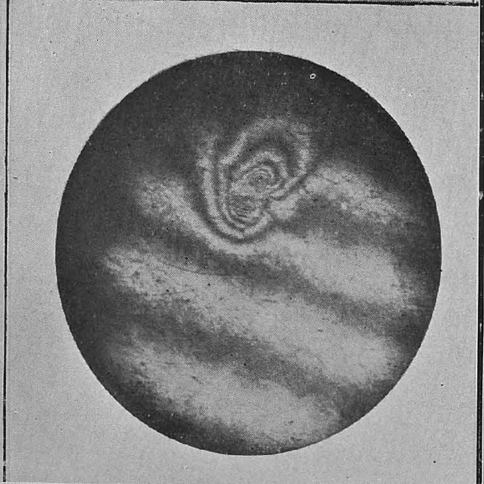


FIG. 6



Percussion Figures: Figs. 1 and 2 in Calcite; Figs. 3 and 4 in Marble; Figs. 5 and 6 in Barytes.

4 JAN 1965

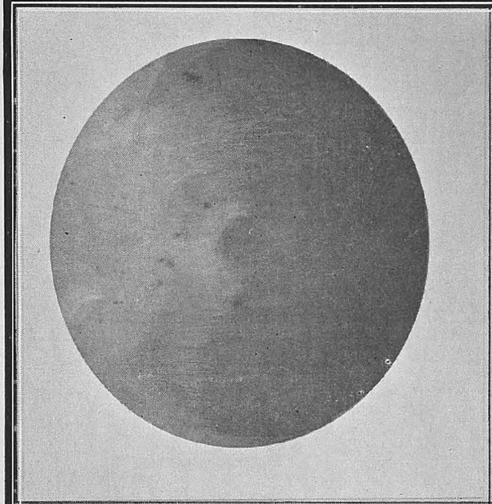


FIG. 1

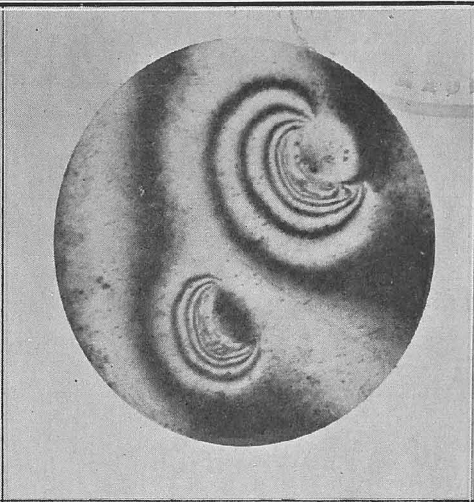


FIG. 2

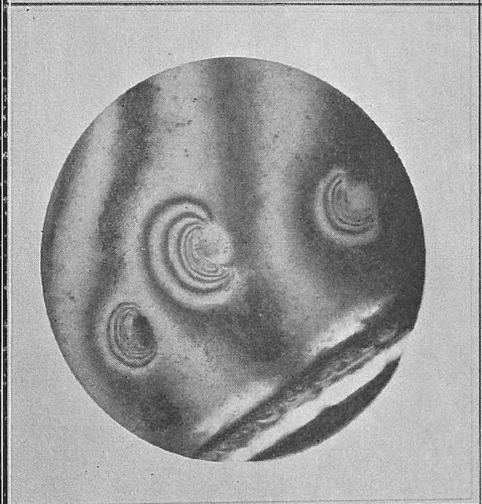


FIG 3

FIG. 4

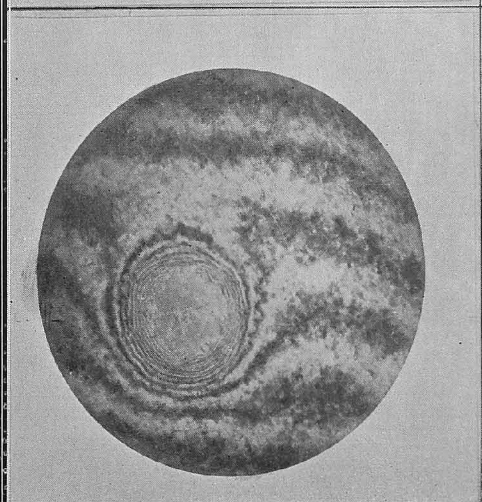


FIG. 5

FIG. 6

Percussion Figures: Figs. 1, 2, 3, 4 in Felspar; Fig. 5, Hard Limestone; Fig. 6, Dolerite.

Reprinted from "The Proceedings of the Indian Academy of Sciences", Vol. XLVIII, No. 6, Sec. A, 1958

THE RAMAN SPECTRA OF CRYSTALLINE SULPHATES OF Ni AND Mn

BY D. KRISHNAMURTI

(*Memoir No.* 112 *from the Raman Research Institute, Bangalore*-6)

Received November 24, 1958

(Communicated by Sir C. V. Raman)

1. INTRODUCTION

BELOW 31·5° C. nickel sulphate crystallizes from solution as $NiSO_4.7H_2O$ in the orthorhombic system. Above this temperature it occurs as crystals of $NiSO_4.6H_2O$ of the tetragonal system and these crystals undergo a transformation to the monoclinic form at 53·3° C. A study of the Raman spectra of these different modifications is obviously of much interest. The Raman spectra of the tetragonal and the orthorhombic varieties of nickel sulphate were reported by the author in an earlier paper; but the results therein were not complete owing to the experimental difficulties met with in recording the spectra of these crystals which are efflorescent. With improved experimental techniques and by the use of larger crystals it has now been possible to obtain fuller information regarding the Raman spectra of nickel sulphate in the tetragonal, orthorhombic and the monoclinic forms. In addition to these results, the present paper reports the details of a study undertaken for the first time on the Raman spectrum of manganese sulphate tetrahydrate.

2. EXPERIMENTAL DETAILS

The crystals of $NiSO_4.7H_2O$, $NiSO_4.6H_2O$ and $MnSO_4.4H_2O$ were grown from aqueous solution by the method of slow evaporation at room temperature. The crystals of $NiSO_4.7H_2O$ were in the form of elongated tablets upto 5 cm. in length. The crystals of $MnSO_4.4H_2O$ occurred as large thick rhombic tablets. Crystals of tetragonal $NiSO_4.6H_2O$ upto a size of 2 cm. square were found to crystallize from solution during the course of particularly warm days of summer. Crystals of the monoclinic $NiSO_4.6H_2O$, however, could not be grown in sufficiently large sizes, although several attempts were made to grow them inside a thermostat at about 60° C. The small crystals grown at this temperature were confirmed to belong to the monoclinic variety from the biaxial figure of small axial angle expected to be exhibited by these crystals under the polarising microscope (Winchell, pp. 225). The tetragonal and orthorhombic crystals of nickel sulphate

were easily identified as such from their external form and depth of colour, and the respective uniaxial and biaxial interference figures exhibited by them served to confirm their identification. The crystals of manganese sulphate were found to be biaxial. They were identified as $MnSO_4.4H_2O$ belonging to the monoclinic system, in view of the close agreement of the observed specific gravity of 2·25 for these crystals with the value of 2·26 given by Winchell.

The Raman spectra were recorded with a medium Hilger quartz spectrograph and using the λ 2536·5 resonance radiation of mercury for the excitation of the Raman effect. A large Littrow quartz spectrograph was also used to record the spectra in the case of $NiSO_4.6H_2O$ and $MnSO_4.4H_2O$. Full details of the experimental set-up have appeared in an earlier paper by the author. All these crystals were efflorescent and were reduced to an opaque white mass after a short exposure to the heat of the quartz arc. This necessitated that the crystals be placed at a safe distance from the arc. In spite of this precaution the crystals deteriorated rapidly and fresh crystals had to be used frequently. In the case of $NiSO_4.7H_2O$ this precaution was particularly necessary since after losing a single water molecule they were transformed into the more stable tetragonal $NiSO_4 \cdot 6H_2O$.

To record the spectrum of $NiSO_4.6H_2O$ above the transformation point, a large tetragonal crystal was chosen and fixed in a brass rod, the lower end of which was immersed in a trough of boiling water and from which steam was rising continuously. Under these conditions the temperature in the neighbourhood of the crystal was about 60° C. It took about half an hour for the crystal to attain the steady temperature of 60° C. and in the process the clear tetragonal crystal was transformed into a translucent polycrystalline mass of the monoclinic variety. Very intense spectra in this case could not be obtained since the parasitic illumination from the polycrystal set a limit to the useful exposure that could be given. While in the other cases exposures of the order of 36 hours were given to obtain intense records of the spectra, exposures of only eight hours could be given in this case.

3. RESULTS AND DISCUSSION

Figs. 2, 3, 4 and 5 in the accompanying Plate reproduce the spectra obtained with $NiSO_4.6H_2O$ (tetragonal), $NiSO_4.6H_2O$ (monoclinic), $NiSO_4.7H_2O$ and $MnSO_4.4H_2O$ respectively. Figs. 1 and 6 of the same Plate show the spectrum of the mercury arc for comparison. Figs. 1, 2, 3 and 4 appearing in the text are the microphotometer records of the spectra with the frequency shifts indicated therein. Table I exhibits the frequency shifts observed in all the four cases.

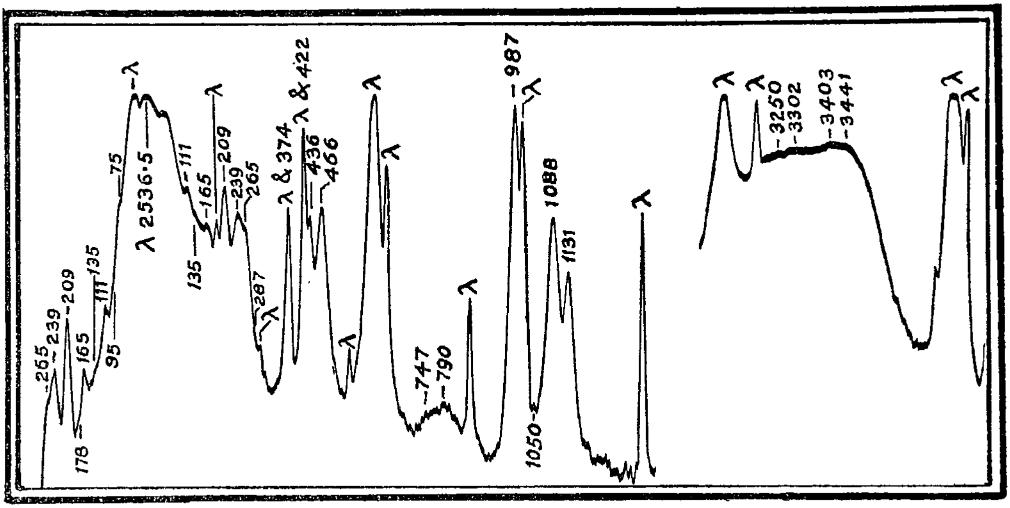


FIG. 1. Microphotometer Record of the Raman Spectrum of $NiSO_4 \cdot 6H_2O$ (Tetragonal).

The group of lines at about 620 cm.$^{-1}$ (ν_3) in the case of $NiSO_4.6H_2O$ (tetragonal) and $NiSO_4.7H_2O$ were observed as anti-Stokes shifts. On the Stokes side only the shifts 620, 620, 622 and 624 could be observed in the four cases. The other components could not be detected owing to the presence in this region of the strong mercury line λ 2576·3 and its satellite λ 2578·4.

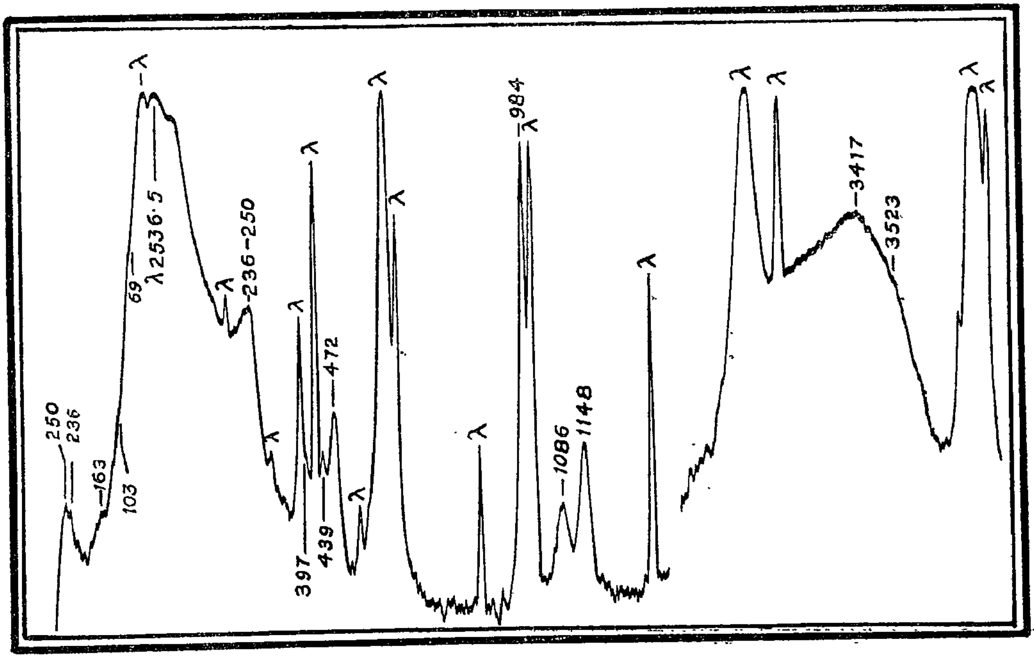


FIG. 2. Microphotometer Record of the Raman Spectrum of $NiSO_4 \cdot 6H_2O$ (Monoclinic).

Similarly the frequency shifts at about 185, 380 and 420 cm.$^{-1}$ shown in Table I were discovered on the anti-Stokes side since on the Stokes side there appear at about these points the mercury lines λ 2548·6, λ 2561·2 and λ 2563·9. The frequency shift at 971 cm.$^{-1}$ close to the intense line 987 cm.$^{-1}$ as well as those at 596 and 620 cm.$^{-1}$ in the case of $NiSO_4.6H_2O$ were manifest in the spectra recorded by the Littrow spectrograph.

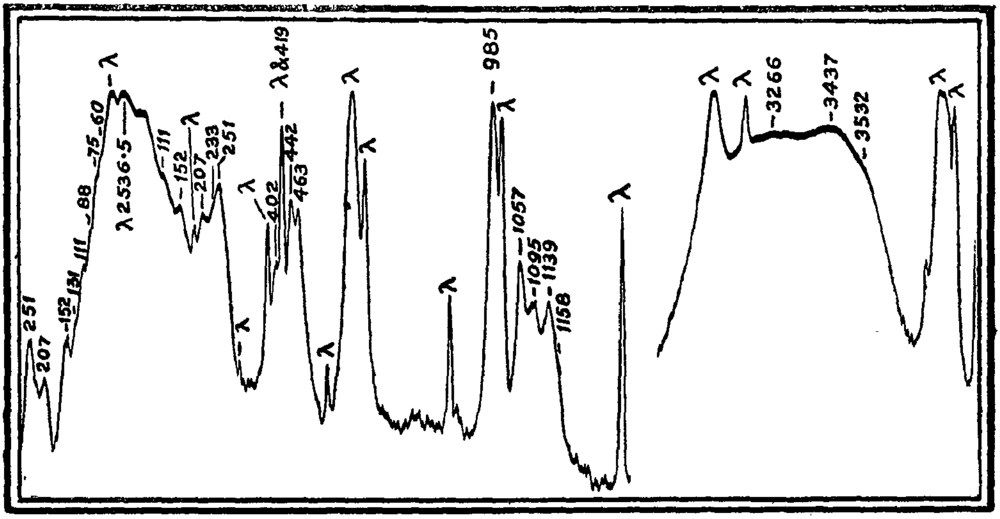


FIG. 3. Microphotometer Record of the Raman Spectrum of $NiSO_4.7H_2O$.

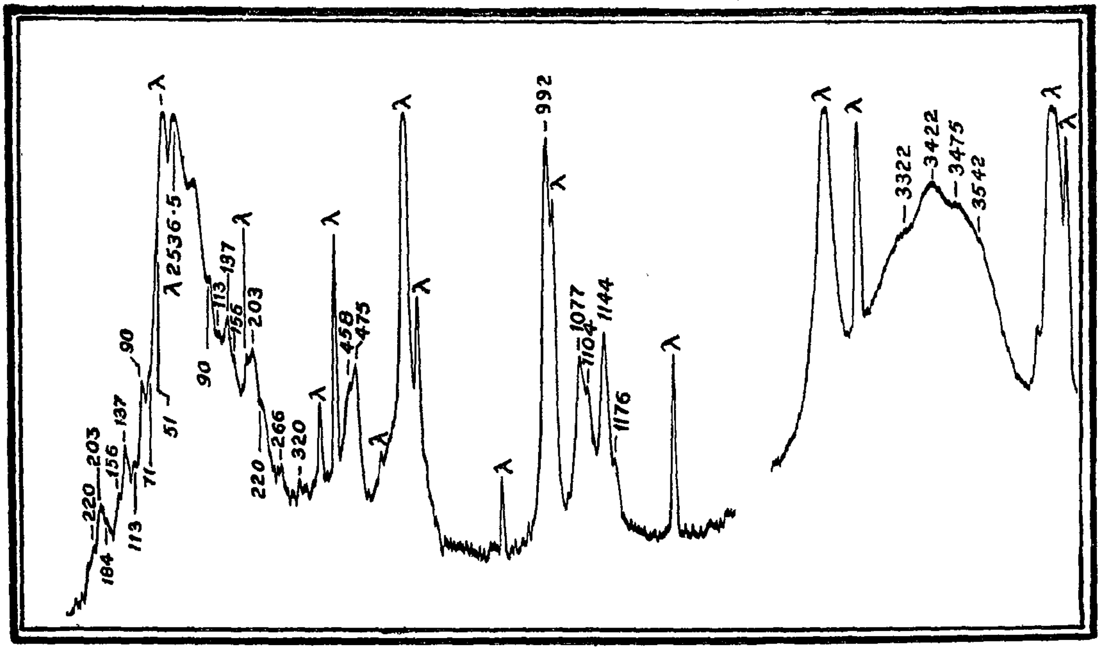


FIG. 4. Microphotometer Record of the Raman Spectrum of $MnSO_4.4H_2O$.

TABLE I

Frequency Shifts in cm.$^{-1}$

	$NiSO_4.6H_2O$ Tetragonal	$NiSO_4.6H_2O$ Monoclinic	$NiSO_4.7H_2O$	$MnSO_4.4H_2O$
Lattice Oscillations	75 s. 95 w. 111 s. 135 v.w. 165 s. 178 v.w. 209 v.s. 239 s. 265 s. 287 v.w.	69 s. 103 w. 163 m. 236 m. 250 s.	60 m. 75 s. 88 m. 111 s. 131 w. 152 s. 207 s. 233 w. 251 v.s.	51 s. 71 m. 90 s. 113 w. 137 s. 156 w. 184 w. 203 s. 220 m. 266 w.d. 320 w.d.
ν_2	374 m. 422 m. 436 s. 466 s.	397 m. 439 m. 472 s.	402 m. 419 w. 442 s. 463 s.	458 m. 475 s.
ν_3	596 m. 620 s. 639 s.	620 m.	612 s. 622 w. 642 m.	624 m.
ν_1	971 w. 987 v.s.	984 v.s.	985 v.s.	992 v.s.
ν_4	1050 v.w. 1088 s. 1131 m.	1086 m. 1148 s.	1057 s. 1095 m. 1139 m. 1158 w.	1077 s. 1104 m. 1144 s. 1176 w.
Water Bands	3250 s. 3302 s. 3403 v.s. 3441 v.s.	3417 v.s. 3523 m.	3266 m. 3437 v.s. 3532 w.	3322 m. 3422 v.s. 3475 s. 3542 m.

s. = Strong; w. = Weak; v.s. = Very Strong; v.w. = Very Weak; m. = Medium; d. = Diffuse.

The observed frequency shifts may readily be seen to fall distinctively into six classes. The frequency shifts upto 320 cm.$^{-1}$ in Table I, all arise due to the rotational and translational oscillations against each other of the ions and water molecules constituting the unit cell of the crystal structure,

Beyond this region we have specific groups of frequency shifts centred at 450, 620, 980 and 1100 cm.$^{-1}$ and which frequency shifts correspond respectively to the (1) doubly degenerate, (2) triply degenerate, (3) totally symmetric and (4) triply degenerate normal modes of vibration of the free SO_4 ion. Lastly we have the broad and diffuse bands characteristic of the water, in the region of frequencies 3200 to 3600 cm.$^{-1}$

In addition to these general features the spectra of the four crystals exhibit several special characteristics of interest. It may be noticed from Table I that in the lattice frequency region striking correspondences exist between the numerical values of the observed frequency shifts. However, an examination of Plate XXX reveals at once that such agreement is not always found to be present with regard to the intensities of the corresponding lines. It may be noticed in particular that the intense frequency shift at 209 cm.$^{-1}$ observed in the tetragonal form of $NiSO_4.6H_2O$ is totally absent in the case of the monoclinic form. In addition, the lattice spectrum in the latter case is smeared out into a diffuse band of illumination with a few definite frequency shifts. Another feature of interest is with regard to the varying multiplicity and relative intensity of the frequency shifts of the SO_4 ions in the different crystals. We shall here seek to explain these features in the light of our knowledge of the crystal structures of these substances.

(1) *$NiSO_4.6H_2O$ (tetragonal)*. The crystal structure of $NiSO_4.6H_2O$ at ordinary temperatures has been determined by Beevers and Lipson (1932) and assigned to the space-group $P4_12_12$ (D_4^4) of the enantiomorphous hemihedral class of the tetragonal system, with four molecules in the unit cell. The nickel and sulphur atoms occupy the special positions of the symmetry of a two-fold axis of rotation. All the other atoms occupy general positions and the six water molecules are octahedrally co-ordinated around the nickel atom. The normal modes of vibration of the system belonging to the symmetry D_4 fall into five representations A_1, A_2, B_1, B_2 and E, A_2 being Raman inactive, and A_1, B_1 and B_2 being infra-red inactive. A study of the infra-red absorption of these crystals would be of value in assigning the frequencies to the different classes.

Since the local symmetry of the SO_4 ion is only C_2 in the crystalline state, the degeneracies are removed and nine frequencies in all (to a first approximation) should manifest themselves in the spectrum. But, owing to the fact that there are four molecules in the unit cell, the number of components is further increased. Group-theoretical analysis shows that in this case the total number of components is twenty-seven, of which twenty-two should be observed in the Raman effect. In all, the spectrum exhibits twelve compo-

nents, some of which are rather broad. The fact of some of the components being too weak, or of their overlapping with each other, might account for the reduced number of observed components. Two broad and weak bands at about 747 and 790 cm.$^{-1}$ are also observed in the spectrum. Their origin is however not quite clear.

(2) *$NiSO_4.7H_2O$ (orthorhombic)*. $NiSO_4.7H_2O$ had been found to belong to the space-group $P2_12_12_1$ ($D_2{}^4$) of the enantiomorphous hemihedral class of the orthorhombic system (Beevers and Schwartz, 1935). All atoms are situated in general positions and the structure is similar to that of the tetragonal hexahydrate with four molecules in the unit cell and with six of the water molecules co-ordinated octahedrally around the nickel atom and the seventh water molecule occupying an interstitial position filling what would otherwise be a hole in the structure. Considerations similar to those put forward in the previous case indicate that due to the internal vibrations of the four SO_4 ions in the unit cell thirty-six components should appear in the Raman effect and twenty-seven in the infra-red. However, apparently for the reasons advanced already, only twelve components—some of which are rather broad—are observed.

The general similarity of the lattice spectrum of $NiSO_4.7H_2O$ (orthorhombic) with that of $NiSO_4.6H_2O$ (tetragonal) has its origin in the similarity of the crystal architecture in these two cases. The intensities of the lines, however, do not correspond, as is naturally to be expected from the fact of the crystals belonging to two different symmetry classes.

(3) *$NiSO_4.6H_2O$ (monoclinic)*. Apart from the fact that the crystals of $NiSO_4.6H_2O$ above 53·3° C. belong to the monoclinic system and exhibit a small optic axial angle of 19° 25′, no data regarding its crystal structure are to be found in the literature. In view of the small optic axial angle observed, it is probable that the monoclinic form of the crystal above 53·3° C. is derived from a distortion of the tetragonal structure observed at ordinary temperatures. In this case seven components corresponding to the internal frequencies of the SO_4 ion are observed. The fact of the doubly degenerate mode of the free ion manifesting itself as three distinct components in the spectrum of the crystal indicates that there are at least two, or more than two molecules in the unit cell of the crystal.

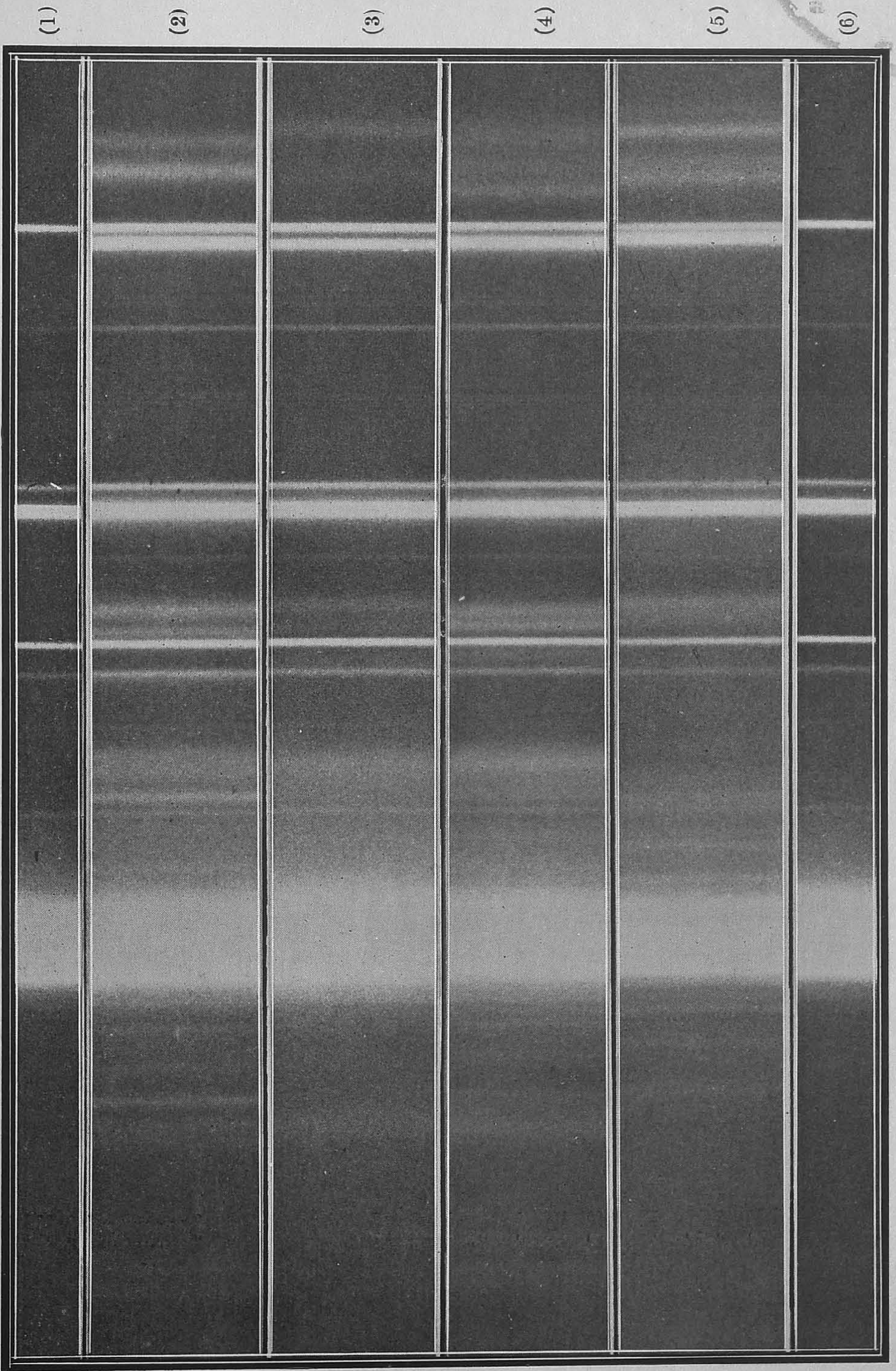
Crystals of $MgSO_4.6H_2O$ exhibit optical characters similar to those of $NiSO_4.6H_2O$ (monoclinic) (*vide* Winchell, pp. 225 and 231). In addition, the values of the angle β are in good agreement, being respectively 81° 26′ and 81° 43′. The structure of $MgSO_4.6H_2O$ has been found to belong to the space-group C2/c ($C_{2h}{}^6$) of the monoclinic system and with eight molecules

in the unit cell (Wyckoff, 1951). The modes of vibration of a system belonging to the group C_{2h} fall into four groups, of which only two groups of symmetric vibrations are Raman active, the other two groups of anti-symmetric vibrations being Raman inactive. If it could be assumed as stated by Sidgwick (1950) that the structure of $NiSO_4.6H_2O$ (monoclinic) is isomorphous with the above $MgSO_4.6H_2O$, we can account for the fact of some of the frequencies like the 209 cm.$^{-1}$ observed in the tetragonal form being totally absent in the monoclinic form. A correspondence could then exist between the vibrations of these two systems and one set of modes present in the tetragonal form would vanish in the monoclinic form. In this respect the numerical agreements between some of the frequencies of the tetragonal form and those of the monoclinic form are also noteworthy. The diffuse nature of the lattice spectrum may have its origin in that the spectrum was recorded at a temperature of 60° C. which is quite close to the transformation temperature of 53·3° C.

(4) *$MnSO_4.4H_2O$*. No data regarding the crystal structure of $MnSO_4.4H_2O$ are available and only its optical characters are known. From the optical data these crystals are known to belong to the monoclinic system. In the case of $MnSO_4.4H_2O$ eight frequency shifts due to the internal vibrations of the SO_4 ions are observed. In particular, the triply degenerate mode, corresponding to the free ion, manifests itself in the crystalline state as four distinct components, thereby indicating that there is more than one molecule in the unit cell of the crystal structure. The lattice spectrum of $MnSO_4.4H_2O$ exhibits close similarities with those observed for $NiSO_4.6H_2O$ (tetragonal) and $NiSO_4.7H_2O$. The atomic weights of nickel and manganese are not very different (being 58·69 and 54·93 respectively) and hence those normal modes which are essentially similar with regard to their rotational or translational characters in the case of these crystals would have nearly equal frequencies. Another noteworthy feature is the comparatively sharp water bands exhibited by $MnSO_4.4H_2O$ unlike those observed in the other cases. The fewer number of water molecules in this case may be the reason for this feature. However, it is not directly possible to correlate the differences observed (in the four cases) with regard to the water bands.

Lakshman Rao (1942) reports for the Raman spectrum of $MnSO_4.7H_2O$ seven lines at 330, 457, 603, 693, 994, 1085 and 1148 cm.$^{-1}$ and two bands due to the water of crystallisation at 3399 and 3467 cm.$^{-1}$ It may be noticed that the frequencies reported by him and those emerging from this study are not in agreement. In particular, the present study does not reveal the existence of lines at 330 and 693 cm.$^{-1}$ even though it is perfectly possible for

4 JAN 1965



Raman Spectra of Crystalline Sulphates. Fig. 2. $NiSO_4 \cdot 6H_2O$ (*tetragonal*). Fig. 3. $NiSO_4 \cdot 6H_2O$ (*monoclinic*). Fig. 4. $NiSO_4 \cdot 7H_2O$. Fig. 5. $MnSO_4 \cdot 4H_2O$. Figs. 1 and 6. Spectrum of the mercury arc.

them to be recorded and observed if present. Lakshman Rao also states that the crystallisations were carried out at about 27° C. It appears that the crystals used by Lakshman Rao were not $MnSO_4.7H_2O$ since $MnSO_4.7H_2O$ is stable only upto 6° C. They could not have been $MnSO_4.4H_2O$ for reasons of disagreement of his results with those of the present study. $MnSO_4$ is known to crystallize with five molecules of water of crystallisation between 6° and 20° C. It is quite probable that his results refer to $MnSO_4.5H_2O$. Apart from the few infra-red frequencies at 635, 1111, 1116, 2940, 3030 and 3125 cm.$^{-1}$ reported by Schaefer and Schubert (1916) from studies on infra-red reflection by $NiSO_4.6H_2O$ (tetragonal) no data are available regarding the infra-red spectra of these crystals for correlation with the Raman frequencies.

In conclusion, the author expresses his sincere thanks to Prof. Sir C. V. Raman, F.R.S., N.L., for his kind interest in this investigation.

SUMMARY

The paper reports the results of investigations on the Raman spectra of $NiSO_4.6H_2O$ (tetragonal and monoclinic forms), $NiSO_4.7H_2O$ (orthorhombic) and $MnSO_4.4H_2O$ (monoclinic). Using λ 2536·5 resonance radiation of mercury for excitation of the Raman effect, it has been possible to observe a large number of frequency shifts. The spectra exhibit striking differences between themselves and also a few characters of general similarity. These are largely explicable in terms of the crystal symmetry and structure of these substances. In all the cases studied, the internal frequencies of the sulphate ion exhibit multiplicities owing to the lowering of the symmetry of the SO_4 ion in the crystalline state and due to the number of ions present in the unit cell.

REFERENCES

1. Beevers and Lipson .. *Z. Kristallogr.*, 1932, **83**, 123.
2. ——— and Schwartz .. *Ibid.*, 1935, **91**, 157.
3. Krishnamurti .. *Proc. Ind. Acad. Sci.*, 1955, **42 A**, 77.
4. ——— .. *Ibid.*, 1957, **46**, 183.
5. Lakshman Rao .. *Ibid.*, 1941, **14**, 48.
6. Schaefer and Schubert .. *Ann. Physik*, 1916, **50**, 283 and 339.
7. Sidgwick .. *Chemical Elements and Their Compounds*, Oxford, 1950, Vol. II, pp. 1437.
8. Winchell .. *Microscopic Characters of Artificial Minerals*, John Wiley & Sons, Inc., N.Y., 1931.
9. Wyckoff .. *Crystal Structures*, Interscience Publishers, Inc., N.Y., 1951, Vol. II, Chap. X, table p. 33.

Reprinted from "The Proceedings of the Indian Academy of Sciences", Vol. L, No. 1, Sec. A, 1959

THE RELATIVISTIC THEORY OF CHEMICAL BINDING

BY K. S. VISWANATHAN, F.A.SC.

(*Memoir No.* 113 *from the Raman Research Institute, Bangalore*-6)

Received March 3, 1959

ABSTRACT

Starting from Breit's relativistic equation for a system of two electrons, it is shown that for a hydrogen molecule (or for a system of two electrons moving in a field of cylindrical symmetry) the component of the total angular momentum (J_z) along the axis of the molecule (axis of symmetry) is a constant of motion. Thus every eigenstate of the system is simultaneously an eigenstate of J_z also, and a state of the system will specify, besides its energy, only the eigenvalue of the component of the angular momentum parallel to the axis of symmetry. The form of the four large components of the wave function relating to their dependence on the azimuthal co-ordinates has been given.

The case of Russel-Saunders approximation has been considered in detail and the nature of the components of the wave function for the singlet and triplet states has been discussed. It is shown that the wave function for the ground state of the hydrogen molecule could be expressed as a sum of a set of symmetric functions of which the first term is the Heitler-London function, and that the wave function for a triplet state should be a superposition of anti-symmetric molecular orbitals. It is shown that relativistic theory brings about in a natural manner the facts relating to the ground state of the molecules C_2 and O_2. Finally, some remarks are made concerning the case of molecules for which the spin-orbit and the spin-spin couplings are strong.

I. INTRODUCTION

THE theoretical solution of problems of molecular chemistry are based usually on one or the other of two methods, known as the Molecular Orbital Theory and the Valence Bond Theory.[1] In the former method, the binding electron is supposed to belong to the entire molecule and is assumed to be moving in the average potential field of the other electrons and the nuclei. The latter is the well-known theory of hydrogen molecule given by Heitler and London. An important consequence of this theory is that the ground state of the hydrogen molecule is the singlet state in which the spins of the two electrons are

opposite to each other and that the states with the spins parallel to each other are unstable. In both the above-mentioned methods, the spin of the electron is introduced as an independent extraneous factor in such a way as to comply with the requirements of the Pauli Exclusion Principle, but is not brought out as an integral part of the theory itself. In view of the very important role that the spin of the electron plays in the chemical binding in molecules, it is desirable to develop a theory based on a relativistic equation which will bring out the spin as an essential part of the mathematical formalism itself, and to analyse the consequences of such a theory. This, precisely, is the object of the present paper.

II. THE HAMILTONIAN OF A SYSTEM OF TWO ELECTRONS

Let us consider a system of two electrons interacting with each other. Let ϕ^{I}, A^{I} and ϕ^{II}, A^{II} denote respectively the scalar and vector potentials of electrons I and II disregarding their interaction. A relativistic wave equation for a system of two electrons interacting with each other and with an external electromagnetic field was given by Breit[2] in the year 1929.

The equation is

$$\left\{ p_0 + \sum_{k=1,2,3} (\alpha_k^{\mathrm{I}} p_k^{\mathrm{I}} + \alpha_k^{\mathrm{II}} p_k^{\mathrm{II}}) + (\alpha_4^{\mathrm{I}} + \alpha_4^{\mathrm{II}}) mc + \frac{e^2}{2cr} \left[\sum_k \alpha_k^{\mathrm{I}} \alpha_k^{\mathrm{II}} + \frac{(\alpha^{\mathrm{I}} \cdot \mathbf{r})(\alpha^{\mathrm{II}} \cdot \mathbf{r})}{r^2} \right] \right\} \psi = 0 \tag{1}$$

where

$$p_k^{\mathrm{I}} = - i\hbar \frac{\partial}{\partial x_k^{\mathrm{I}}} + \left(\frac{e}{c}\right) \mathrm{A}_k^{\mathrm{I}};$$

$$p_k^{\mathrm{II}} = - i\hbar \frac{\partial}{\partial x_k^{\mathrm{II}}} + \left(\frac{e}{c}\right) \mathrm{A}_k^{\mathrm{II}}$$

and

$$p_0 = \frac{i\hbar}{c} \frac{\partial}{\partial t} + \frac{e}{c} (\phi^{\mathrm{I}} + \phi^{\mathrm{II}}) - \frac{e^2}{cr} \tag{2}$$

The matrices α_k^{I} and α_k^{II} are the usual Dirac matrices, but now they operate on the first and second suffixes respectively. The representations that we choose for these are given in Appendix I. The wave function ψ has sixteen components which will be designated by ψ_{mn} $(m, n = 1, 2, 3, 4)$, the first

suffix m referring to electron I and the second suffix n referring to electron II. If a^{I} and b^{II} be any of the above-mentioned matrices, then

$$(a^{\text{I}}b^{\text{II}}\psi)_{m,n} = \sum_{k,l} a_{mk}{}^{\text{I}} b_{nl}{}^{\text{II}} \psi_{kl} \tag{3}$$

Equation (1) is thus equivalent to sixteen differential equations in the sixteen functions ψ_{mn}.

By writing in (1)

$$\alpha_1 = \rho_1\sigma_1; \qquad \alpha_2 = \rho_1\sigma_2; \qquad \alpha_3 = \rho_1\sigma_3 \tag{4}$$

where ρ_1, σ_1, σ_2, σ_3 are the matrices (4) of Appendix I, we see that the Hamiltonian of the system is given by

$$\begin{aligned} \mathrm{H} = e\,(\phi^{\text{I}} + \phi^{\text{II}}) &+ \frac{e^2}{r} - \rho_1{}^{\text{I}}c\,(\sigma^{\text{I}} \cdot \mathbf{p}^{\text{I}}) \\ &- \rho_1{}^{\text{II}}c\,(\sigma^{\text{II}} \cdot \mathbf{p}^{\text{II}}) - mc^2\,(\alpha_4{}^{\text{I}} + \alpha_4{}^{\text{II}}) \\ &- \frac{e^2}{2r}\left[(\rho_1{}^{\text{I}}\sigma^{\text{I}} \cdot \rho_1{}^{\text{II}}\sigma^{\text{II}}) + \frac{(\rho_1{}^{\text{I}}\sigma^{\text{I}} \cdot \mathbf{r})\,(\rho_1{}^{\text{II}}\sigma^{\text{II}} \cdot \mathbf{r})}{r^2}\right] \end{aligned} \tag{5}$$

III. The x-Component of the Total Angular Momentum

We shall consider the case of the hydrogen molecule which is the simplest example of a two-electron molecular system. Let the origin be chosen at one of the nuclei and let the line of action of the molecule coincide with the positive direction of the x-axis. If the nuclei are denoted by the suffixes A and B, the potentials ϕ^{I} and ϕ^{II} are given by

$$\phi^{\text{I}} = e\left(\frac{1}{r_{a_1}} + \frac{1}{r_{b_1}}\right)$$

and

$$\phi^{\text{II}} = e\left(\frac{1}{r_{a_2}} + \frac{1}{r_{b_2}}\right)$$

where $r_a = \sqrt{x^2 + y^2 + z^2}$ and $r_b = \sqrt{(x - \mathrm{R})^2 + y^2 + z^2}$, R being the internuclear distance.

Let $m_x{}^{\text{I}}$ and $m_x{}^{\text{II}}$ denote the x-component of the orbital angular momenta of electrons I and II respectively. Then the dynamical variable

$M_x = (m_x^{I} + m_x^{II})$ represents the x-component of the total orbital angular momentum of the system. In the same way, the quantity

$$S_x = \frac{\hbar}{2}(\sigma_x^{I} + \sigma_x^{II})$$

gives the component of the total spin angular momentum parallel to the axis of the molecule. The sum J_x of these two (*i.e.*)

$$J_x = m_x^{I} + m_x^{II} + \frac{\hbar}{2}(\sigma_x^{I} + \sigma_x^{II}) \tag{6}$$

represents therefore the x-component of the *total* angular momentum of the system parallel to the axis of the molecule.

It has been proved in Appendix II that

$$J_x H - H J_x = 0 \tag{7}$$

One can see from the Appendix that neither S_x nor M_x commutes with the Hamiltonian but only their sum has the property of commuting with H.

Thus the component of the total angular momentum parallel to the axis of the molecule is a constant of motion.

Since J_x commutes with H, every eigenfunction of H will be simultaneously an eigenfunction of J_x also. An eigenstate, thus, is specified by two parameters, namely, the energy of the state and the eigenvalue of the component of the total angular momentum parallel to the axis of the molecule.

IV. THE EIGEN ψ'S OF J_x

Let us denote the eigenvalues of J_x by $M\hbar$ where M is yet undetermined. The equation

$$J_x \psi = M\hbar\psi \tag{8 a}$$

or

$$\left[m_x^{I} + m_x^{II} + \frac{\hbar}{2}(\sigma_x^{II} + \sigma_x^{II})\right]\psi = M\hbar\psi \tag{8 b}$$

will then be equivalent to sixteen equations in the sixteen components of ψ. We write here only four of these equations, namely, those involving the four large components ψ_{33}, ψ_{34}, ψ_{43} and ψ_{44} of ψ. From (3), it follows that the (3, 3), (3, 4), (4, 3) and (4, 4)th components of the equation (8) are given by

$$-i\left(\frac{\partial}{\partial\phi_1}+\frac{\partial}{\partial\phi_2}\right)\psi_{33}=(M-1)\,\psi_{33}$$

$$-i\left(\frac{\partial}{\partial\phi_1}+\frac{\partial}{\partial\phi_2}\right)\psi_{34}=M\psi_{34}$$

$$-i\left(\frac{\partial}{\partial\phi_1}+\frac{\partial}{\partial\phi_2}\right)\psi_{43}=M\psi_{43}$$

$$-i\left(\frac{\partial}{\partial\phi_1}+\frac{\partial}{\partial\phi_2}\right)\psi_{44}=(M+1)\,\psi_{44} \qquad (9)$$

Hence

$$\psi_{33}=f_{33}(r_1, r_2, \theta_1, \theta_2)\, e^{i\,(\overline{m_1-1}\phi_1+m_2\phi_2)}$$

$$\psi_{34}=f_{34}(r_1, r_2, \theta_1, \theta_2)\, e^{i\,(m_1\phi_1+m_2\phi_2)}$$

$$\psi_{43}=f_{43}(r_1, r_2, \theta_1, \theta_2)\, e^{i\,(m_1\phi_1+m_2\phi_2)}$$

$$\psi_{44}=f_{44}(r_1, r_2, \theta_1, \theta_2)\, e^{i\,(m_1\phi_1+\overline{m_2+1}\phi_2)} \qquad (10)$$

where $m_1+m_2=M$.

Since the ψ's should be single-valued, the numbers m_1 and m_2 are integers.

One can verify that the other twelve small components satisfy differential equations of the type (9) and their form is therefore similar to the expressions given by (10).

Now it has been shown by Breit[2, 3] that equation (1) involving sixteen components can be reduced to an approximate equation with four components expressed in terms of Pauli's spin matrices. This reduced equation is given by (6) of Appendix I and is equivalent to a set of four equations in the four large components of ψ which, following Breit, we denote by ψ_{33}, ψ_{34}, ψ_{43} and ψ_{44} respectively. Since the equation (6) of Appendix I is sufficient for all physical purposes and gives quantitative results correct to the order of $(v/c)^2$, we shall in the following sections be discussing the consequences following from this only.

The matrices σ^{I} and σ^{II} in (6) of Appendix I are the Pauli matrices and they operate on the suffixes 3 and 4. The operation rule for the product of any two matrices on ψ is still governed by (3), but now the summation for k or l extends over the numbers 3 and 4 only. One can verify that the equation $J_x\psi=M\hbar\psi$ is identical with the four equations given by (9), and the form of the four large components is therefore still given by the equations (10).

By solving the simultaneous differential equations (7), Appendix I, one can determine the functions ψ_{33}, ψ_{34}, ψ_{43} and ψ_{44} respectively. Thus a state ψ which we denote by the column vector $[\psi_{33}, \psi_{34}, \psi_{43}, \psi_{44}]$ consists in general of four non-zero components.

V. THE RUSSEL-SAUNDERS COUPLING

When the spin-orbit interaction is small compared to the electrostatic interaction of the two electrons, the terms involving B^{I}, B^{II}, C_{xy} as well as the last four terms in the expression for A in equation (6, Appendix I) could be ignored in the first approximation. This is the assumption underlying the case of the well-known Russel-Saunders coupling. The equations (7, Appendix I) now reduce to

$$(H - W)\,\psi_{33} = 0; \qquad (H - W)\,\psi_{34} = 0$$
$$(H - W)\,\psi_{43} = 0; \qquad (H - W)\,\psi_{44} = 0 \tag{11}$$

where H is the non-relativistic Hamiltonian of the system. The above equations are satisfied by four different states of the system in each of which all components, excepting one, are zero. Thus, solutions of (11) exist which have the form

$$\psi_1 = [\psi_{33}, 0\ 0, 0]; \qquad \psi_2' = [0, \psi_{34}, 0, 0]$$
$$\psi_3' = [0, 0, \psi_{43}, 0]; \qquad \psi_4 = [0, 0, 0, \psi_{44}]$$

Since the spin-orbit coupling terms are neglected in the Hamiltonian (6, Appendix I), one can see that both

$$M_x = m_x^{I} + m_x^{II} \qquad \text{and} \qquad S_x = \frac{\hbar}{2}(\sigma_x^{I} + \sigma_x^{II})$$

are constants of motion in this case, and further ψ_1 and ψ_4 are states belonging to the eigenvalues $\hbar$ and $-\hbar$ of S_x respectively, whereas the states ψ_2' and ψ_3' have zero as the eigenvalue of S_x. From the form of ψ_{34} and ψ_{43} in (10), one can see that the states

$$\psi_2 = \psi_2' + \psi_3' = [0, \psi_{34}, \psi_{43}, 0]$$

and

$$\psi_3 = \psi_2' - \psi_3' = [0, \psi_{34}, -\psi_{43}, 0]$$

have the eigenvalue M for the x-component of the total orbital angular momentum and zero for the x-component of the total spin angular momentum. Further if

$$\Sigma^2 = \tfrac{1}{4}\left[(\sigma_x^{I} + \sigma_x^{II})^2 + (\sigma_y^{I} + \sigma_y^{II})^2 + (\sigma_z^{I} + \sigma_z^{II})^2\right]$$

one can verify from (3) that the states ψ_1, ψ_2 and ψ_4 are eigenstates of Σ^2 belonging to the eigenvalue 2 (or the value 1 for Σ), and the state ψ_3 has the eigenvalue zero for Σ. The former set of states are thus the well-known triplet states of two electron systems whereas the state ψ_3 represents the singlet state.

Now according to the Pauli Exclusion Principle, every state should be anti-symmetric with respect to an exchange of the two electrons. An exchange operator, which we denote by P, interchanges the spin states of the two electrons as well as their positions and thus under its influence, a component of the state $\psi_{mn}(1, 2)$ gets transformed into $\psi_{nm}(2, 1)$. The state

$$\psi = [\psi_{33}(1, 2), \psi_{34}(1, 2), \psi_{43}(1, 2), \psi_{44}(1, 2)]$$

therefore changes into $\psi = [\psi_{33}(2, 1), \psi_{43}(2, 1), \psi_{34}(2, 1), \psi_{44}(2, 1)]$ under the influence of P. In order to satisfy the Exclusion Principle, the following relations should therefore be obeyed by the components of a state: (1) $\psi_{33}(2, 1) = -\psi_{33}(1, 2)$; (2) $\psi_{34}(2, 1) = -\psi_{43}(1, 2)$ and (3) $\psi_{44}(2, 1) = -\psi_{44}(1, 2)$. One can see that the states ψ_1 and ψ_4 will be anti-symmetric if the components ψ_{33} and ψ_{44} are anti-symmetric functions of the position of the two electrons. Since ψ_{34} and ψ_{43} are solutions of the same Schroedinger equation in the Russel-Saunders approximation, the state ψ_2 will be anti-symmetric if its components ψ_{34} and ψ_{43} are such that $\psi_{34}(1, 2) = \psi_{43}(1, 2) = -\psi_{34}(2, 1) = -\psi_{43}(2, 1)$, and similarly the state ψ_3 will be anti-symmetric if

$$\psi_{34}(1, 2) = \psi_{34}(2, 1) = -\psi_{43}(1, 2) = -\psi_{43}(2, 1)$$

(*a*) *The symmetric and anti-symmetric states.*—Let $\psi_{nlm}(r_{a_1})$ ...... denote a complete set of eigenfunctions of a hydrogen atom whose nucleus is at position A. Then these functions form a complete basis for the expansion of any arbitrary single-valued function of the position $\mathbf{r}_1 = (x_1, y_1, z_1)$ of the electron 1. Similarly, th set of eigenfunctions ($\psi_{n'l'm'}$ $(\mathbf{r}_{b_2})$ of a hydrogen atom with its nucleus in position B form a complete set of eigenfunctions for the expansion of an arbitrary function of the position $\mathbf{r}_2$ of the electron 2. Then, by a well-known theorem,[4] the set of functions

$$\psi_{nlm}(r_{a_1})\,\psi_{n'l'm'}(r_{b_2}) \begin{pmatrix} n = 1, 2 \,\ldots\ldots\ldots \\ n' = 1, 2 \,\ldots\ldots\ldots \end{pmatrix}$$

form a complete basis for the expansion of an arbitrary function $\psi(\mathbf{r}_1, \mathbf{r}_2)$ of the two electrons, that vanishes when $r_1 \to \infty$ and $r_2 \to \infty$. Thus we can write

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \sum_{nn'} \mathrm{F}_{nn'}\,\psi_n(\mathbf{r}_{a_1})\,\psi_{n'}(\mathbf{r}_{b_2}) \tag{12}$$

where we have written n and n' for the set of three quantum numbers (n, l, m) and (n', l', m') of the hydrogen atoms.

Now if the wave function $\psi(\mathbf{r}_1, \mathbf{r}_2)$ is symmetric with respect to an interchange of the two electrons, we get from (12) by interchanging the suffixes (1) and (2) that

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \sum_{nn'} \mathrm{F}_{nn'} \psi_n(\mathbf{r}_{a_2}) \psi_{n'}(\mathbf{r}_{b_1}) \tag{13}$$

The above equation alternatively follows from the expansion of $\psi(\mathbf{r}_1, \mathbf{r}_2)$ in terms of the set of functions $\psi_n(\mathbf{r}_{a_2}) \psi_{n'}(\mathbf{r}_{b_1}) \ldots$ which also form a complete basis for the expansion of a function of the two electrons.

By adding (12) and (13)* we see that any symmetric function of the electrons 1 and 2 can be written as

$$\psi^+(\mathbf{r}_1, \mathbf{r}_2) = \sum_{nn'} f_{nn'} [\psi_n(\mathbf{r}_{a_1}) \psi_{n'}(\mathbf{r}_{b_2}) + \psi_n(\mathbf{r}_{a_2}) \psi_{n'}(\mathbf{r}_{b_1}) \tag{14}$$

where $f_{nn'} = \frac{1}{2} \mathrm{F}_{nn'}$. The suffix + has been used to indicate that the function ψ is symmetric with respect to an interchange of the electrons.

In the same way we can see that an anti-symmetric function $\psi^-(\mathbf{r}_1, \mathbf{r}_2)$ of the electrons has the form

$$\psi^-(r_1, r_2) = \sum_{nn'} g_{nn'} [\psi_n(\mathbf{r}_{a_1}) \psi_{n'}(\mathbf{r}_{b_2}) - \psi_n(\mathbf{r}_{a_2}) \psi_{n'}(\mathbf{r}_{b_1})] \tag{15}$$

(*b*) *The singlet state.*—We have seen before that the singlet state ψ_3 is represented by the vector $[0, \psi_{34}^+, -\psi_{34}^+, 0]$ where ψ_{34}^+ is a symmetric function of the two electrons. Now from the equations (10) it follows that ψ_{34} has the form given by

$$\psi_{34} = f_{34}(r_1, r_2, \theta_1, \theta_2) e^{i(m_1\phi_1 + m_2\phi_2)} \tag{10 a}$$

where m_1 and m_2 are integers satisfying the relation $m_1 + m_2 = \mathrm{M}$. Thus the general form of ψ_{34} consists of a superposition of terms of the type (10 *a*) for all integral values of m_1 and m_2 satisfying the condition $m_1 + m_2 = \mathrm{M}$.

Let us consider the case where the component of the total angular momentum parallel to the axis of the molecule is zero, *i.e.*, $m_1 + m_2 = 0$ or $m_1 = -m_2$. It is well known that the ground state of the hydrogen molecule belongs to this species. Thus from (14) and (10 *a*) it follows that the eigenfunction of the ground state is given by

$$\psi_{34}^+(\mathbf{r}_1, \mathbf{r}_2) = \sum_{nlm} \sum_{n'l'} f_{nn'} [\psi_{nlm}(\mathbf{r}_{a_1}) \psi_{n'l'-m}(\mathbf{r}_{b_2}) + \psi_{nlm}(\mathbf{r}_{a_2}) \psi_{n'l'-m}(\mathbf{r}_{b_1})] \tag{16}$$

* We assume the absolute convergence of these series.

The first term of the above series is

$$[\psi_{100}(\mathbf{r}_{a_1})\,\psi_{100}(\mathbf{r}_{b_2}) + \psi_{100}(\mathbf{r}_{a_2})\,\psi_{100}(\mathbf{r}_{b_1})] \tag{17}$$

which is the well-known Heitler-London wave function. This function explains qualitatively in a satisfactory manner the chemical binding of the hydrogen molecule, but quantitatively it gives a value of 3·14 e.v. for the binding energy of the molecule, the observed value being 4·72 e.v. The quantitative inaccuracy of the Heitler-London function is not surprising when we recall that it is only the first term in the expansion of the wave function in terms of symmetric wave functions built from the hydrogen atomic orbitals. Since the discrepancy between the energy value derived from (17) and the experimentally observed one is not very large, one can hope that the series (16) converges rather rapidly and that the addition of a few more terms to (17) will lead one to the correct value for the binding energy of the molecule.

(*c*) *The triplet states.*—As an example of the triplet states, consider the state ψ_1 which is described by a non-vanishing anti-symmetric component $\psi_{33}^-(\mathbf{r}_1, \mathbf{r}_2)$ and three vanishing components. From (10) we have

$$\psi_{33} = f_{33}(r_1, r_2, \theta_1, \theta_2)\, e^{i(m_1\phi_1 + m_2\phi_2)} \tag{18}$$

where $m_1 + m_2 = (M - 1)$.

Since the component of the total spin angular momentum for this state is $\hbar$, the component of the total orbital angular momentum is equal to $(M - 1)\hbar$. The general form of ψ_{33} will consist of a superposition of terms of the type (18) for all integral values of m_1 and m_2 satisfying the relation $m_1 + m_2 = (M - 1)$. Thus from (15) it follows that ψ_{33}^- is given by

$$\psi_{33}^-(\mathbf{r}_1, \mathbf{r}_2) = \sum_{nlm} \sum_{n'l'm'} g_{nn'} [\psi_{nlm}(\mathbf{r}_{a_1})\,\psi_{n'l'm'}(\mathbf{r}_{b_2}) - \psi_{nlm}(\mathbf{r}_{a_2})\,\psi_{n'l'm'}(\mathbf{r}_{b_1}) \tag{19}$$

where $m + m' = (M - 1)$.

In other words, the state ψ_1 is a superposition of all anti-symmetric molecular states for which the component of the total orbital angular momentum is equal to $(M - 1)\hbar$. In the same way we can obtain the wave function for the state ψ_4 by superposing anti-symmetric molecular states have the eigenvalue $(M + 1)\hbar$ for the component of the total orbital angular momentum parallel to the axis of the molecule.

Regarding the state ψ_2 we shall consider the particular case $M = 0$. ψ_2 is described by the vector $[0, \psi_{34}^-, \psi_{34}^-, 0]$ where ψ_{34}^- is anti-symmetric

with respect to an interchange of the two electrons. From (15) and (10 a) it follows that the form of ψ_{34}^- is given by

$$\psi_{34}^-(\mathbf{r}_1, \mathbf{r}_2) = \sum_{nlm} \sum_{n'l'} g_{nn'} [\psi_{nlm}(\mathbf{r}_{a_1}) \psi_{n'l'-m}(\mathbf{r}_{b_2}) - \psi_{nlm}(\mathbf{r}_{a_2}) \psi_{n'l'-m}(\mathbf{r}_{b_1}) \quad (20)$$

The first term of the above series is

$$[\psi_{100}(\mathbf{r}_{a_1}) \psi_{100}({}_b\mathbf{r}_2) - \psi_{100}({}_a\mathbf{r}_2) \psi_{100}(\mathbf{r}_{b_1})] \quad (21)$$

It is a well-known fact that this function leads to an unstable chemical state.

(*d*) *The case of* C_2 *and* O_2.—From (19) or (20), it follows that the wave function for a triplet state is a superposition of a number of terms each of which is anti-symmetric with respect to an interchange of the two electrons. Further, all the terms in a series belong to the same species, *i.e.*, the states belonging to these terms have all the same eigenvalue for the x-component of the total orbital angular momentum.

According to the Heitler-London theory, the ground state of the hydrogen molecule is the singlet state and the triplet states are unstable. The theory therefore requires as an essential condition for a stable bond that the spins of the two electrons should be antiparallel. This rule is obeyed by a large class of molecules, but nevertheless, exceptions are also present of which we may cite the exampl(s of C_2 and O_2. The ground state of C_2 is the state designated in spectroscopic notation by $^3\Pi_u$ whereas for oxygen, the ground state is $^3\Sigma_g^-$. Oxygen is in addition paramagnetic. Heitler[6] and Nordheim-Poschl[7] showed that by adding to the wave function (21) a few more terms of the same species, or in other words, by taking into account the interaction between states of the same species, the deviations between theory and experiment can be removed. Now in the Russel-Saunders approximation, Σ has the value 1 for the triplet states and Σ_z has the values 1, 0, $-$ 1 respectively for the states ψ_1, ψ_2 and ψ_4. Further each term in (19) or (20) corresponds to the same value of the component of the total orbital angular momentum parallel to the axis of the molecule. The expansion (19) is, therefore, nothing but a superposition of states of the same species. Relativistic theory offers thus a clear explanation of the facts relating to C_2 and O_2 and brings about naturally the concept of interaction of states of the same species.

VI. STRONG SPIN-ORBIT COUPLING

The discussion of the last section holds good only for the case of Russel-Saunders coupling. Strictly speaking, the Russel-Saunders coupling and the

j-j coupling are two extreme ideal cases and the coupling in most cases is intermediate. When the spin-orbit interaction is strong compared to the electrostatic interaction of the two electrons, the above discussion will break down. In such a case, only the component of the total angular momentum along the axis is a constant of motion and none of the quantities Σ, Σ_x and M_x is a constant of motion. An eigenstate of the energy will be an eigenstate of J_x also and conversely every state of the system will specify, besides the energy, only the x-component of the total angular momentum. For the case of strong spin-orbit coupling, the usual spectroscopic notation which presumes that all the quantities Σ, Σ_x and M_x are constants of motion, is therefore inconsistent with theoretical considerations and requires modification.

When the spin-orbit coupling is strong, the solution of the equations (7, Appendix I) will lead to non-zero values for all the four components of the wave function. Further it follows from equations (6) and (7) of Appendix I that a state will be anti-symmetric with respect to an interchange of electrons only if the components have the following properties:

(1) $\psi_{33}(2, 1) = -\psi_{33}(1, 2)$;

(2) $\psi_{34}(2, 1) = -\psi_{43}(1, 2)$;

(3) $\psi_{44}(2, 1) = -\psi_{44}(1, 2)$.

Thus ψ_{33} and ψ_{44} are anti-symmetric functions, while ψ_{34} or ψ_{43} can be a mixture of both symmetric and anti-symmetric functions. The form of the former components is therefore given by (15) while the latter can be expanded as a series of the type (12). Now the energy of the system is given by

$$W = \frac{\int \psi^* H \psi d\tau_1 d\tau_2}{\int \psi^* \psi d\tau_1 d\tau_2} \qquad (22)$$

where ψ^* is the transpose of the vector ψ and is given by the *row* vector $[\psi_{33}^*, \psi_{34}^*, \psi_{43}^*, \psi_{44}^*]$.

In practice, the computation of the energy will be facilitated if (1) the series for the function ψ_{33}, etc., are terminated at a particular point and their coefficients determined by the variational principle and (2) if use is made of the tables giving the values of exchange, coulomb and other related integrals.[8, 9, 10]

We can now interpret the quantities ψ_{33}^2, $\psi_{34}^2 + \psi_{43}^2$ and ψ_{44}^2 as the probabilities of finding the spins of the electrons respectively (1) parallel to the x-axis, (2) opposite to each other and (3) parallel to the negative direction of the x-axis. The ground state of the molecule will be the state having the

least value for the energy of all the states, subject to the condition that it is stable besides.

The author's grateful thanks are due to Professor Sir C. V. Raman for his kind interest in this work and valuable suggestions.

REFERENCES

1. Coulson, C. A. .. *Valence* (Oxford), 1952.
2. Breit, G. .. *Phys. Rev.*, 1929, **34**, 553.
3. ——— .. *Ibid.*, 1932, **39**, 616.
4. Courant, R. and Hilbert, D. .. *Methods of Mathematical Physics* (Interscience Publishers, Inc., N.Y.), 1953, p. 56.
5. Heitler, W. and London, F. .. *Zeit. f. Phys.*, 1927, **44**, 455.
6. ——— .. *Marx's Hand. d. Radiol.*, 1934, **2**, 485.
7. Nordheim-Poschl, G. .. *Ann. Phys.*, 1936, **26**, 258.
8. Kotani, M., Ishigaro, E., Hijikata, K. and others *J. Phy. Soc. of Japan*, 1953, **8**, 463.
9. Brennan R. O. and Mulligan, J. F. *Jour. Ch. Phy.*, 1952, **20**, 1635.
10. Mullikan, R. S., Rieke, C. A., Orloff, D. and Orloff, H. *Ibid.*, 1949, **17**, 1248.

Appendix I

For the sake of easy reference, we collect here some of the results and equations of Breit that are made use of in the paper. The relativistic wave equation of Breit is given by

$$\left[\begin{array}{c} p_0 + \sum\limits_{k=1,2,3} (\alpha_k^{I} p_k^{I} + \alpha_k^{II} p_k^{II}) + (\alpha_4^{I} + \alpha_4^{II})\, mc \\ + \dfrac{e^2}{2cr}\left\{ \sum\limits_k \alpha_k^{I}\alpha_k^{II} + \dfrac{(\alpha^{I}\cdot \mathbf{r})(\alpha^{II}\cdot \mathbf{r})}{r^2} \right\} \end{array}\right] \psi = 0 \qquad (1)$$

Here p_0, p^{I} and p^{II} are given by equations (2) of Section II. The wave function ψ has sixteen components which are denoted by ψ_{mn} ($m, n = 1, 2, 3, 4$). The matrices α^{I} and α^{II} are the usual Dirac matrices, but now they operate on the suffixes I and II respectively. They are given by

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}; \quad \alpha_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix};$$

$$\alpha_3 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}; \quad \alpha_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad (2)$$

If we set

$$\alpha_1 = \rho_1\sigma_1; \quad \alpha_2 = \rho_1\sigma_2; \quad \alpha_3 = \rho_1\sigma_3, \qquad (3)$$

where

$$\sigma_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix};$$

$$\sigma_3 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}; \quad \rho_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

equation (1) becomes

$$\left\{ \begin{aligned} p_0 + \rho_1{}^{\mathrm{I}} (\sigma^{\mathrm{I}} \cdot \mathbf{p}^{\mathrm{I}}) + \rho_1{}^{\mathrm{II}} (\sigma^{\mathrm{II}} \cdot \mathbf{p}^{\mathrm{II}}) + (\alpha_4{}^{\mathrm{I}} + \alpha_4{}^{\mathrm{II}})\, mc \\ + \frac{e^2}{2cr} \left[(\rho_1{}^{\mathrm{I}}\sigma^{\mathrm{I}} \cdot \rho_1{}^{\mathrm{II}}\sigma^{\mathrm{II}}) + \frac{(\rho_1{}^{\mathrm{I}}\alpha^{\mathrm{I}} \cdot \mathbf{r})\,(\rho_1{}^{\mathrm{II}}\alpha^{\mathrm{II}} \cdot r)}{r^2} \right. \end{aligned} \right\} \psi = 0 \qquad (5)$$

It follows that the Hamiltonian of the system is given by equation (5) of Section II.

Next the Breit equation in the Schroedinger-Pauli form is given by

$$\begin{aligned} \Big\{ \mathrm{E} - 2mc^2 + e\mathrm{V} - \frac{1}{2m} [(p^{\mathrm{I}})^2 + (p^{\mathrm{II}})^2] + \frac{(p^{\mathrm{I}})^4 + (p^{\mathrm{II}})^4}{8m^3c^3} \\ + \frac{e^2}{2m^2c^2} [r^{-1} (\mathbf{p}^{\mathrm{I}}, \mathbf{p}^{\mathrm{II}}) + r^{-3} (x_i{}^{\mathrm{II}} - x_i{}^{\mathrm{I}})(x_j{}^{\mathrm{II}} - x_j{}^{\mathrm{I}})\, p_i{}^{\mathrm{I}} p_j{}^{\mathrm{II}}] \\ + \frac{eh}{4\pi mc} [(\mathbf{H}^{\mathrm{I}}, \sigma^{\mathrm{I}}) + (\mathbf{H}^{\mathrm{II}}, \sigma^{\mathrm{II}})] + \frac{ehi}{8\pi m^2c^2} [(\mathscr{E}^{\mathrm{I}}, \mathbf{p}^{\mathrm{I}}) + (\mathscr{E}^{\mathrm{II}}, \mathbf{p}^{\mathrm{II}})] \\ + \frac{eh}{8\pi mc^2} \Big[\Big(\mathscr{E}^{\mathrm{I}} \times \frac{\mathbf{p}^{\mathrm{I}}}{m}, \sigma^{\mathrm{I}}\Big) + \Big(\mathscr{E}^{\mathrm{II}} \times \frac{\mathbf{p}^{\mathrm{II}}}{m}, \sigma^{\mathrm{II}}\Big) \\ + 2e\Big(\frac{\mathbf{r}^{\mathrm{I}} - \mathbf{r}^{\mathrm{II}}}{r^3} \times \frac{\mathbf{p}^{\mathrm{II}}}{m}, \sigma^{\mathrm{I}}\Big) + 2e\Big(\frac{\mathbf{r}^{\mathrm{II}} - \mathbf{r}^{\mathrm{I}}}{r^3} \times \frac{\mathbf{p}^{\mathrm{I}}}{m}, \sigma^{\mathrm{II}}\Big)\Big] \\ - \Big(\frac{eh}{4\pi mc}\Big)^2 \langle (\nabla^{\mathrm{II}}\sigma^{\mathrm{I}})(\nabla^{\mathrm{I}}\sigma^{\mathrm{II}})\, r^{-1} \rangle \Big\} \psi = 0 \end{aligned} \qquad (6)$$

where

$$\mathbf{H}^{\mathrm{I}} = \operatorname{curl} \mathbf{A}^{\mathrm{I}}, \quad \mathbf{H}^{\mathrm{II}} = \operatorname{curl} \mathbf{A}^{\mathrm{II}}$$

$$\mathscr{E}^{\mathrm{I}} = -\operatorname{grad}^{\mathrm{I}} \mathrm{V} = -\Big(\frac{\partial}{\partial x^{\mathrm{I}}}, \frac{\partial}{\partial y^{\mathrm{I}}}, \frac{\partial}{\partial z^{\mathrm{I}}}\Big) \mathrm{V};$$

$$\mathrm{V} = \phi^{\mathrm{I}} + \phi^{\mathrm{II}} - \frac{e}{r}.$$

$\langle\,\rangle$ means the differential operators $\mathbf{p}^{\mathrm{I}}$, $\mathbf{p}^{\mathrm{II}}$ apply only within the $\langle\,\rangle$. The σ's are the Pauli matrices and have the representations given by

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Let us write

$$\mathrm{E} - 2mc^2 = \mathrm{W};$$

$$\mathrm{A} = e\mathrm{V} - \frac{1}{2m}\{(p^{\mathrm{I}})^2 + (p^{\mathrm{II}})^2\} + \frac{(p^{\mathrm{I}})^4 + (p^{\mathrm{II}})^4}{8m^3c^3}$$

$$+ \frac{e^2}{2m^2c^2}\,[r^{-1}\,\mathbf{p}^{\mathrm{I}} \cdot \mathbf{p}^{\mathrm{II}} + \Sigma\, r^{-3}\,(x_i^{\mathrm{II}} - x_i^{\mathrm{I}})\,(x_j^{\mathrm{II}} - x_j^{\mathrm{I}})\,p_i^{\mathrm{I}} p_j^{\mathrm{II}}]$$

$$+ \frac{ehi}{8\pi m^2c^2}\,\{(\mathcal{E}^{\mathrm{I}}, \mathbf{p}^{\mathrm{I}}) + (\mathcal{E}^{\mathrm{II}}, \mathbf{p}^{\mathrm{II}})\}.$$

$$\mathrm{B}^{\mathrm{I}} = \frac{eh}{8\pi mc^2}\left[\mathcal{E}^{\mathrm{I}} \times \frac{\mathbf{p}^{\mathrm{I}}}{m} + 2e\,\frac{\mathbf{r}^{\mathrm{I}} - \mathbf{r}^{\mathrm{II}}}{r^3} \times \frac{p^{\mathrm{II}}}{m}\right] = (\mathrm{B}_x{}^{\mathrm{I}}, \mathrm{B}_y{}^{\mathrm{I}}, \mathrm{B}_z{}^{\mathrm{I}})$$

$$\mathrm{B}^{\mathrm{II}} = \frac{eh}{8\pi mc^2}\left[\mathcal{E}^{\mathrm{II}} \times \frac{p^{\mathrm{II}}}{m} + 2e\,\frac{\mathbf{r}^{\mathrm{II}} - \mathbf{r}^{\mathrm{I}}}{r^3} \times \frac{\mathbf{p}^{\mathrm{I}}}{m}\right] = (\mathrm{B}_x{}^{\mathrm{II}}, \mathrm{B}_y{}^{\mathrm{II}}, \mathrm{B}_z{}^{\mathrm{II}})$$

and

$$-\left(\frac{eh}{4\pi mc}\right)^2 \langle(\nabla^{\mathrm{II}}, \sigma^{\mathrm{I}})\,(\nabla^{\mathrm{I}}, \sigma^{\mathrm{II}})\,r^{-1}\rangle = \Sigma\,\mathrm{C}_{xy}\sigma_x{}^{\mathrm{I}}\sigma_y{}^{\mathrm{II}}$$

where

$$\mathrm{C}_{xy} = -\left(\frac{eh}{4\pi mc}\right)^2 \frac{\partial}{\partial x^{\mathrm{II}}}\,\frac{\partial}{\partial y^{\mathrm{I}}}\,r^{-1}.$$

Then the equation (6) is equivalent to the following four equations:

$$(\mathrm{A} + \mathrm{B}_x{}^{\mathrm{I}} + \mathrm{B}_x{}^{\mathrm{II}} + \mathrm{C}_{xx} + \mathrm{W})\,\psi_{33} + (\mathrm{B}_y{}^{\mathrm{II}} - i\mathrm{B}_z{}^{\mathrm{II}} + \mathrm{C}_{xy} - i\mathrm{C}_{xz})\,\psi_{34}$$
$$+ (\mathrm{B}_y{}^{\mathrm{I}} - i\mathrm{B}_z{}^{\mathrm{I}} + \mathrm{C}_{yx} - i\mathrm{C}_{zx})\,\psi_{43} + (\mathrm{C}_{yy} - \mathrm{C}_{zz}$$
$$- i\mathrm{C}_{yz} + \mathrm{C}_{zy})\,\psi_{44} = 0.$$

$$(\mathrm{B}_y{}^{\mathrm{II}} + i\mathrm{B}_z{}^{\mathrm{II}} + \mathrm{C}_{xy} + i\mathrm{C}_{xz})\,\psi_{33} + (\mathrm{A} + \mathrm{B}_x{}^{\mathrm{I}} - \mathrm{B}_x{}^{\mathrm{II}} - \mathrm{C}_{xx} + \mathrm{W})\,\psi_{34}$$
$$+ (\mathrm{C}_{yy} + \mathrm{C}_{zz})\,\psi_{43} + (\mathrm{B}_y{}^{\mathrm{I}} - i\mathrm{B}_z{}^{\mathrm{I}} - \mathrm{C}_{yx} + i\mathrm{C}_{zx})\,\psi_{44} = 0 \tag{7}$$

$$(\mathrm{B}_y{}^{\mathrm{I}} + i\mathrm{B}_z{}^{\mathrm{I}} + \mathrm{C}_{yx} + i\mathrm{C}_{zx})\,\psi_{33} + (\mathrm{C}_{yy} + \mathrm{C}_{zz})\,\psi_{34}$$
$$+ (\mathrm{A} + \mathrm{B}_x{}^{\mathrm{II}} - \mathrm{B}_x{}^{\mathrm{I}} - \mathrm{C}_{xx} + \mathrm{W})\,\psi_{43} + (\mathrm{B}_y{}^{\mathrm{II}} - i\mathrm{B}_z{}^{\mathrm{II}} - \mathrm{C}_{xy}$$
$$+ i\mathrm{C}_{xz})\,\psi_{44} = 0$$

$$(C_{yy} - C_{zz} + i\overline{C_{yz} + C_{zy}})\,\psi_{33} + (B_y^{I} + iB_z^{I} - C_{yx} - iC_{zx})\,\psi_{34} + (B_y^{II} + iB_z^{II} - C_{xy} - iC_{xz})\,\psi_{43} + (A - B_x^{I} - B_x^{II} + C_{xx} + W)\,\psi_{44} = 0.$$

APPENDIX II

Proof of the relation $J_xH - HJ_x = 0$.—We have seen that the Hamiltonian for a hydrogen molecule is given by

$$H = -e(\phi^{I} + \phi^{II}) + \frac{e^2}{r} - \rho_1^{I}c(\sigma^{I}\cdot \mathbf{p}^{I}) - \rho_1^{II}c(\sigma^{II}\cdot \mathbf{p}^{II}) - m_0c^2(a_4^{I} + a_4^{II}) - \frac{e^2}{2r}\left\{\rho^{I}\sigma^{I}\cdot \rho_1^{II}\sigma^{II}) + \frac{(\rho_1\sigma^{I}\cdot \mathbf{r})\ (\rho_1^{II}\sigma^{II}\cdot \mathbf{r})}{r^2}\right\} \tag{1}$$

where

$$\phi^{I} = e\left(\frac{1}{r_{a_1}} + \frac{1}{r_{b_1}}\right) \quad \text{and} \quad \phi^{II} = e\left(\frac{1}{r_{a_2}} + \frac{1}{r_{b_2}}\right);$$

r is the interelectronic distance. Each of the angular momentum variables m_x^{I} and m_x^{II} commutes with r_{a_1}, r_{b_1}, r_{a_2} and r_{b_2} and hence with the potentials ϕ^{I} and ϕ^{II}. Further

$$\begin{aligned}[m_x^{I} + m_x^{II}, r^2) &= [m_x^{I} + m_x^{II}, \Sigma x_1^2 + \Sigma x_2^2 - 2\Sigma x_1x_2] \\ &= -2(y_1z_2 - y_2z_1) - 2(y_2z_1 - y_1z_2) = 0\end{aligned} \tag{2}$$

$(m_x^{I} + m_x^{II})$ commutes with r^2 and hence with any function of r^2, particularly with $1/r$ and $1/r^3$.

Now

$$i\hbar(\dot{m}_x^{I} + \dot{m}_x^{II}) = (m_x^{I} + m_x^{II})H - H(m_x^{I} + m_x^{II}) \tag{3}$$

Since $(m_x^{I} + m_x^{II})$ commutes with ϕ^{I}, ϕ^{II}, $1/r$ and $1/r^3$, we have

$$\begin{aligned}i\hbar(\dot{m}_x^{I} + \dot{m}_x^{II}) = &- c\rho_1^{I}\{m_x^{I}(\sigma^{I}\cdot \mathbf{p}^{I}) - (\sigma^{I}\cdot p^{I})\,m_x^{I}\} \\ &- c\rho_1^{II}\{m_x^{II}(\sigma^{II}\cdot p^{II}) - (\sigma^{II}\cdot p^{II})\,m_x^{II}\} \\ &- \frac{e^2}{2r^3}\begin{Bmatrix}(m_x^{I} + m_x^{II})(\rho_1^{I}\sigma^{I}\cdot \mathbf{r})(\rho_1^{II}\sigma^{II}\cdot \mathbf{r}) \\ -(\rho_1^{I}\sigma^{I}\cdot \mathbf{r})\ (\rho_1^{II}\sigma^{II}\cdot \mathbf{r})(m_x^{I} + m_x^{II})\end{Bmatrix}\end{aligned} \tag{4}$$

Now

$$\begin{aligned}\{m_x^{I}(\sigma^{I}\cdot \mathbf{p}^{I}) - (\sigma^{I}\cdot \mathbf{p}^{I})\,m_x^{I}\} &= (\sigma^{I}, m_x^{I}\mathbf{p}^{I} - \mathbf{p}^{I}m_x^{I}) \\ &= i\hbar(\sigma_2p_3 - \sigma_3p_2)^{I}\end{aligned} \tag{5}$$

Similarly

$$\{m_x^{\mathrm{II}}(\sigma^{\mathrm{II}} \cdot p^{\mathrm{II}}) - (\sigma^{\mathrm{II}} \cdot p^{\mathrm{II}})\, m_x^{\mathrm{II}}\} = i\hbar\,(\sigma_2 p_3 - \sigma_3 p_2)^{\mathrm{II}} \tag{6}$$

Next the following Poisson bracket relations can be verified:

$$\left[m_x^{\mathrm{I}} + m_x^{\mathrm{II}},\ \frac{(x_2 - x_1)(y_2 - y_1)}{r^3}\right] = \frac{(x_2 - x_1)(z_2 - z_1)}{r^3}\ ;$$

$$\left[m_x^{\mathrm{I}} + m_x^{\mathrm{II}},\ \frac{(x_2 - x_1)(z_2 - z_1)}{r^3}\right] = \frac{(x_2 - x_1)(y_2 - y_1)}{r^3}\ ;$$

$$\left[m_x^{\mathrm{I}} + m_x^{\mathrm{II}},\ \frac{(y_2 - y_1)(z_2 - z_1)}{r^3}\right] = \left[\frac{[(z_2 - z_1)^2 - (y_2 - y_1)^2}{r^3}\right]\ ;$$

$$\left[m_x^{\mathrm{I}} + m_x^{\mathrm{II}},\ \frac{(y_2 - y_1)^2}{r^3}\right] = \frac{2\,(y_2 - y_1)(z_2 - z_1)}{r^3}\ ;$$

$$\left[m_x^{\mathrm{I}} + m_x^{\mathrm{II}},\ \frac{(z_2 - z_1)^2}{r^3}\right] = \frac{-2\,(y_2 - y_1)(z_2 - z_1)}{r^3}\,. \tag{7}$$

Hence if we denote the components of the vectors $\rho_1^{\mathrm{I}}\sigma^{\mathrm{I}}$ and $\rho_1^{\mathrm{II}}\sigma^{\mathrm{II}}$ by

$$(a_x^{\mathrm{I}}, a_y^{\mathrm{I}}, a_z^{\mathrm{I}}) \qquad \text{and} \qquad (a_x^{\mathrm{II}}, a_y^{\mathrm{II}}, a_z^{\mathrm{II}}),$$

we get from (3), (4), (5), (6) and (7) that

$$\begin{aligned} &i\hbar\,(\dot m_x^{\mathrm{I}} + \dot m_x^{\mathrm{II}}) \\ &\quad = -\,i\hbar\, c\rho_1^{\mathrm{I}}(\sigma_2 p_3 - \sigma_3 p_2)^{\mathrm{I}} - i\hbar\, c\rho_1^{\mathrm{II}}(\sigma_2 p_3 - \sigma_3 p_2)^{\mathrm{II}} \\ &\qquad \frac{-\,i\hbar e^2}{2r^3}\left\{\begin{array}{l} (a_x^{\mathrm{I}}a_y^{\mathrm{II}} + a_y^{\mathrm{I}}a_x^{\mathrm{II}})(x_2 - x_1)(z_2 - z_1) \\ \quad -(a_x^{\mathrm{I}}a_z^{\mathrm{II}} + a_z^{\mathrm{I}}a_x^{\mathrm{II}})(x_2 - x_1)(y_2 - y_1) \\ \quad +(a_y^{\mathrm{I}}a_z^{\mathrm{II}} + a_z^{\mathrm{I}}a_y^{\mathrm{II}})[(z_2 - z_1)^2 - (y_2 - y_1)^2] \\ \quad +2\,(a_y^{\mathrm{I}}a_y^{\mathrm{II}} - a_z^{\mathrm{I}}a_z^{\mathrm{II}})(y_2 - y_1)(z_2 - z_1) \end{array}\right\} \end{aligned} \tag{8}$$

Next we have

$$\begin{aligned} (\sigma_x^{\mathrm{I}}\sigma^{\mathrm{I}} - \sigma^{\mathrm{I}}\sigma_x^{\mathrm{I}},\ \mathbf{p}^{\mathrm{I}}) &= 2i\,(\sigma_3 p_2 - \sigma_2 p_3)^{\mathrm{I}} \\ (\sigma_x^{\mathrm{II}}\sigma^{\mathrm{II}} - \sigma^{\mathrm{II}}\sigma_x^{\mathrm{II}},\ \mathbf{p}^{\mathrm{II}}) &= 2i\,(\sigma_3 p_2 - \sigma_2 p_3)^{\mathrm{II}} \end{aligned} \tag{9}$$

Also

$$\left[\sigma_x^{\mathrm{I}} + \sigma_x^{\mathrm{II}},\ \frac{(\rho_1^{\mathrm{I}}\sigma_1^{\mathrm{I}} \cdot \rho_1^{\mathrm{II}}\sigma^{\mathrm{II}})}{r^3}\right] = 0 \tag{10}$$

Further

$$\sigma_x^{\mathrm{I}}(\rho_1^{\mathrm{I}}\sigma^{\mathrm{I}}\cdot\mathbf{r})(\rho_1^{\mathrm{II}}\sigma^{\mathrm{II}}\cdot\mathbf{r})-(\rho_1^{\mathrm{I}}\sigma^{\mathrm{I}}\cdot\mathbf{r})(\rho_1^{\mathrm{II}}\sigma^{\mathrm{II}}\cdot\mathbf{r})\sigma_x^{\mathrm{I}}$$
$$=2i\{a_z^{\mathrm{I}}(y_2-y_1)-a_y^{\mathrm{I}}(z_2-z_1)\}\{a_x^{\mathrm{II}}(x_2-x_1)+a_y^{\mathrm{II}}(y_2-y_1)+a_z^{\mathrm{II}}(z_2-z_1)\}$$

$$\sigma_x^{\mathrm{II}}(\rho_1^{\mathrm{I}}\sigma^{\mathrm{I}}\cdot\mathbf{r})(\rho_1^{\mathrm{II}}\sigma^{\mathrm{II}}\cdot\mathbf{r})-(\rho_1^{\mathrm{I}}\sigma^{\mathrm{I}}\cdot\mathbf{r})(\rho_1^{\mathrm{II}}\sigma^{\mathrm{II}}\cdot\mathbf{r})\sigma_x^{\mathrm{II}}$$
$$=2i\{a_z^{\mathrm{II}}(y_2-y_1)-a_y^{\mathrm{II}}(z_2-z_1)\}\{a_x^{\mathrm{I}}(x_2-x_1)+a_y^{\mathrm{I}}(y_2-y_1)+a_z^{\mathrm{I}}(z_2-z_1)\} \tag{11}$$

From (9), (10) and (11), we get

$$i\hbar(\dot{\sigma}_x^{\mathrm{I}}+\dot{\sigma}_x^{\mathrm{II}})$$
$$=(\sigma_x^{\mathrm{I}}+\sigma_x^{\mathrm{II}})\mathrm{H}-\mathrm{H}(\sigma_x^{\mathrm{I}}+\sigma_x^{\mathrm{II}})$$
$$=-2ic\rho_1^{\mathrm{I}}(\sigma_3p_2-\sigma_2p_3)^{\mathrm{I}}-2ic\rho_1^{\mathrm{II}}(\sigma_3p_2-\sigma_2p_3)$$
$$-\frac{2ie^2}{2r^3}\left\{\begin{aligned}&(a_x^{\mathrm{I}}a_z^{\mathrm{II}}+a_z^{\mathrm{I}}a_x^{\mathrm{II}})(y_2-y_1)(x_2-x_1)\\&+(a_z^{\mathrm{I}}a_y^{\mathrm{II}}+a_z^{\mathrm{II}}a_y^{\mathrm{II}})\{(y_2-y_1)^2-(z_2-z_1)^2\}\\&+2(a_z^{\mathrm{I}}a_z^{\mathrm{II}}-a_y^{\mathrm{I}}a_y^{\mathrm{II}})(y_2-y_1)(z_2-z_1)\\&-(a_y^{\mathrm{I}}a_x^{\mathrm{II}}+a_y^{\mathrm{II}}a_x^{\mathrm{I}})(z_2-z_1)(x_2-x_1)\end{aligned}\right\} \tag{12}$$

From (8) and (12) it follows that

$$\dot{m}_x^{\mathrm{I}}+\dot{m}_x^{\mathrm{II}}+\frac{\hbar}{2}(\dot{\sigma}_x^{\mathrm{I}}+\dot{\sigma}_x^{\mathrm{II}})=0.$$

Or, in other words,

$$\mathrm{J}_x=m_x^{\mathrm{I}}+m_x^{\mathrm{II}}+\frac{\hbar}{2}(\sigma_x^{\mathrm{I}}+\sigma_x^{\mathrm{II}})\text{ is a constant of motion.}$$

Fields possessing cylindrical symmetry.—Though the above result has been proved for the case of a hydrogen molecule only, it holds good in general for any system of two electrons moving in a field possessing cylindrical symmetry. If the x-axis is chosen as the axis of symmetry, the potentials ϕ^{I} and ϕ^{II} in this case are functions of two variables only, *i.e.*, $\phi^{\mathrm{I}}=\phi^{\mathrm{I}}(x_1, r_1)$ and $\phi^{\mathrm{II}}=\phi^{\mathrm{II}}(x_2, r_2)$, r_1 and r_2 being the distances of the electrons from the origin. Since $\mathrm{M}_x=m_x^{\mathrm{I}}+m_x^{\mathrm{II}}$ commutes with all the variables x_1, x_2, r_1 and r_2, it commutes with the first term in the Hamiltonian (1). The remaining terms do not involve ϕ^{I} or ϕ^{II} and are the same as for the hydrogen molecule. Thus for a system of two electrons moving in a field of cylindrical symmetry, the component of the total angular momentum along the axis of symmetry is a constant of motion. Since for most diatomic molecules and individual bonds in complex molecules the charge cloud approximately satisfies cylindrical symmetry, the above result can be applied to such cases also.

Reprinted from "The Proceedings of the Indian Academy of Sciences", Vol. XLIX, No. 4, Sec. A, 1959

CHRISTIAAN HUYGHENS AND THE WAVE THEORY OF LIGHT

BY SIR C. V. RAMAN

(*Memoir No.* 114 *from the Raman Research Institute, Bangalore–6*)

Received March 20, 1959

1. INTRODUCTION

THE world, as we perceive it, is pictured for us by the rays of light which proceed from each point of the objects around us and form images of these objects on the retinæ of our eyes—a statement which needs no amendment even when the aid of optical instruments such as the telescope or the microscope is invoked to enlarge our powers of vision. To put it a little differently, the principles of geometrical optics suffice to describe the behaviour of light as commonly experienced, *viz.*, that the rays of light are propagated in straight lines; that the angles of incidence and of reflection are equal; and that in refraction the rays of light are bent according to the law of sines. In his celebrated *Traite de la Lumiere* published in the year 1690, Christiaan Huyghens showed that these facts of experience are consistent with the hypothesis that light is in the nature of wave-motion propagated through space and can indeed be satisfactorily explained on the basis of that hypothesis. The treatise of Huyghens contains much other material of importance; a perusal of it leaves on the mind of the reader the impression that it is a masterpiece of scientific thought and exposition which possesses an enduring value and interest.

In connection with some experimental investigations on the diffraction of light undertaken by the present author—the results of which will soon be published in these *Proceedings*—the need was felt for a careful study of the original ideas of Huyghens. The task was made much easier by using the literal translation of his treatise from the original French into English by Sylvanus P. Thompson published by the Chicago University Press in the year 1912. The results of the study were surprising; it emerged that the ideas of Huyghens were not fully or even correctly understood by later writers. This misunderstanding has had some far-reaching consequences. Especially in regard to the so-called " Principle of Huyghens " do we find that later writers have chosen a path for which there is no warrant in the writings of Huyghens. In view of these circumstances, it has appeared desirable to put forward a clear exposition of the ideas of Huyghens and supplement the same

by a critical examination of the writings of later authors which claim to be based on those ideas.

2. The Nature of Light

In the first few pages of his book Huyghens set out the considerations which led him to infer that light is in the nature of a movement which spreads into space in all directions from a luminous source. He remarks that the terrestrial sources which are observed to emit light, such as fire or flame evidently contain bodies in rapid motion. Then again, when sunlight is collected by a concave mirror and concentrated on material objects, it has the same effects as fire, *viz.*, it disunites the particles of those objects. It is therefore natural to suppose that light is itself some kind of motion and that the sensation of light is excited when such movement is communicated to the nerves at the back of the human eye. Huyghens also remarks on the extreme speed with which light spreads on every side and on the circumstance that when light comes from different regions, even from those directly opposite each other, the rays traverse one another without hindrance. The facts indicate that light is a movement transmitted through space and not a transport of matter which reaches the eye from the source of light.

To account for the very high velocity of propagation of light—known from the observations of Römer on the eclipses of Jupiter's satellites—and the fact that light can pass through empty space, Huyghens proposed a physical picture of the ætherial medium which could explain its power to transmit waves with such high velocity. He suggested that the æther of space consists of an immense number of extremely small and extremely hard spherical particles in close contact with each other. Experiments on the percussion of elastic solid spheres on each other show that a medium of the nature postulated could propagate waves simultaneously in all directions with high velocity and in such manner that waves travelling in different directions at one and the same time would not hinder each other's progress. Huyghens further recognized that every luminous object would necessarily contain an immense number of centres emitting light and that from each of such centres thousands of waves might emerge in the smallest imaginable time; he pointed out that these considerations would make it easier to understand, why in spite of the enfeeblement of the individual waves by their spread through immense distances, the light of the distant stars continues to be perceptible to human eyes.

3. The Rectilinear Propagation of Light

The mechanical model of the æther proposed by Huyghens to account for the propagation of waves of light through it also enabled him to give a

simple and satisfactory explanation of why light travels out from the original source in straight lines. Considering the medium which transmits light to be composed of an immense number of very small and very hard spherical particles in close contact with each other, it follows that each of these particles when it is displaced from its position by the passage over it of the parent wave sent out from the original source would itself function as the source of a wave which spreads out from it in all directions. The particular or partial waves of this nature recognized by Huyghens in his argument would be as numerous as the total number of the individual particles of æther contained in a sphere drawn with the source of the primary wave as centre and the distance to which it has travelled out as the radius. It is evident, also, that these partial waves, though present in enormous numbers within the volume of the sphere, would individually be of excessively feeble force. In theory, they are assumed to travel out from their respective centres in all directions but actually, by reason of their excessive feebleness the effect of all the partial waves may be totally ignored except of those which arrive together at the same instant, in other words, *simultaneously*, at the point of observation. For, in the latter case, their effects would be superposed and would add up to give an observable result. Such simultaneity in arrival and consequent superposition of effects would be possible only in the case of those partial waves which originate at points in the medium between the primary source of light and the point of observation which lie on the straight line joining them. The summation of the partial waves originating at such points and reaching the point of observation at the same instant would produce the luminous effect there observed. Thus the light which reaches the point of observation from the original source may be considered as having travelled out along the straight line joining the two points.

4. The Reflection and Refraction of Light

In the second and third chapters of his treatise, Huyghens considered the phenomena which arise when light travelling in one medium reaches the boundary which separates it from another. In considering those phenomena, various physical questions arise and are discussed by him, as for example, why some materials are transparent to light and others are opaque, why the velocity of light in a material medium should differ from that in empty space and what the actual configuration of the boundary of the separation between two media is or can be.

The explanation of reflection and refraction in terms of the wave hypothesis is based on an idea which Huyghens found himself compelled to introduce, *viz.*, *that the elements of the area* of the boundary of separation between

the two media act as sources of partial waves which travel out respectively into the two media. The partial waves returned from the boundary into the first medium build up the regularly reflected wave, while those travelling into the second medium give the refracted wave.

A simple geometric construction based on the foregoing ideas enabled Huyghens to explain the familiar law of reflection of light. Likewise, assuming the second medium to be transparent and that the velocity of light in it differs from that in the first, a similar geometric construction led him to the result that the sine of the angle of incidence and the sine of the angle of refraction bear to each other a constant ratio which is the same as the ratio of the velocities of light in the two media. The phenomena of the total internal reflection of light also found a satisfactory explanation. Huyghens showed further that his construction leads directly to Fermat's principle of minimum time for the passage of light from one point to another when the two points are in different media.

5. THE WAVE-OPTICS OF HUYGHENS

Before proceeding to comment on the writing of later authors on the work of Huyghens, we may usefully here summarise the basic concepts of his theory. Huyghens put forward and sought to establish the proposition that when a wave of light diverges from its source, every small portion of the wave is capable of propagating itself independently with the same velocity as the rest of it; in an isotropic medium, the direction of such propagation is the wave-normal and hence this is also the direction of the ray in the sense of geometrical optics. The same idea forms the basis of Huyghens' explanation of the reflection and refraction of light. When the elements of area of an advancing wave-front reach the boundary between two media, each such element gives rise, respectively in the two media, to the elements of area in the reflected and refracted waves. These latter advance normally to themselves in such a direction that they can join up and form continuous wave-fronts. The geometric constructions employed by Huyghens enable these requirements to be satisfied. The propagation of light in an inhomogeneous medium considered in the fourth chapter of Huyghens' treatise can also be very simply dealt with on the same basis. The elements of area of the wave-front in such a medium advance normally to themselves with the velocity appropriate to their positions in the medium. As they advance, they join up to form new wave-fronts which are orthogonal to the path of the light-rays in the medium.

Later writers have criticised the arguments employed by Huyghens in his treatise. One remark which is often made is that the theory of Huyghens

would result in his wave-fronts moving backwards as well as forwards and that he had given no explanation for the absence of backward propagation. But this criticism is not justified and is itself based on a misunderstanding. Huyghens was concerned with the behaviour of an *advancing* wave-front in a homogeneous medium. The partial waves which in his theory give the observed light intensity by their superposition are those which diverge from points lying on the straight line between the source and the observer; in order to reach the observer simultaneously they should all move *away* from the source and *towards* the point of observation, in other words move *forwards* towards the observer. The possibility of backward propagation is thus ruled out completely.

Another criticism which has frequently been advanced is that the theory of Huyghens is based on an arbitrary assumption, *viz.*, that only along the envelope of his partial waves would there by any observable intensity of light. This criticism is also based on a misunderstanding. It should be remembered that Huyghens was unaware that the waves of light are periodic disturbances having a definite wave-length. He assumed that light consists of *individual* waves which diverge in all directions from the original source and the partial waves contemplated in his theory would therefore also be of the same nature. The build-up of a finite intensity from the superposition of a very large number of such waves, each of which is extremely feeble, would accordingly be possible only if they arrive *simultaneously* at the point of observation. The diagram appearing in the first chapter of Huyghens' treatise is intended to assist the reader to appreciate the arguments set out in the text; *viz.*, at *each* point on the wave-front a great number of partial waves arrive *simultaneously* and build up the intensity at that point, while the entire wave may be itself considered as made up of a great number of elementary areas at which the light-intensity has thus been built up. In the later chapters in which Huyghens' theories of reflection and refraction and of the propagation of light in an inhomogeneous medium are expounded, the diagrams are intended to exhibit how the complete wave-front arising from these processes is built up out of its elementary parts or areas. Here again, the final result is an individual wave, and it may therefore be correctly described as the envelope of the partial waves which co-operate in building it up.

6. The Partial Waves of Huyghens

Since the concept of partial waves introduced by Huyghens in his treatise has played an important role in physical optics, it is appropriate that we consider it here in some detail. Though the words appear in several chapters

of his treatise, it should be remarked that they do not have the same significance in each case. In the first chapter which seeks to explain the rectilinear propagation of light, the partial waves arise as a consequence of the assumed discrete structure of the luminiferous medium; each particle in the medium is regarded as a source of such waves. In the second and third chapters, the partial waves are assumed to arise when the primary wave reaches the boundary separating the two media with different properties. The elements of area of the boundary are here regarded as the source of partial waves. Since they travel with different velocities, they are distinct from each other in the two media. In the fourth chapter which deals with the propagation of light in inhomogeneous media, the partial waves are assumed to diverge from the elements of area of the advancing wave-front in such a medium.

If the luminiferous medium were empty space, the assumption that it consists of discrete particles which can function as emitters of partial waves would be difficult to justify. In the case of material media, however, there is good reason for assuming that the discrete atoms of which they are composed could function as sources of secondary or partial waves. Even so, however, these partial waves would reinforce each other in the direction of propagation of the primary wave and merge with it, while in other directions they would interfere and cancel out each other's effects. Thus, they would, in all cases, cease to be observable. Accordingly, the notion of partial waves can, in such circumstances, be regarded only as hypothetical or virtual and not as an observable or physical reality. The same remarks would also be applicable in regard to the propagation of light in a medium which is inhomogeneous. Indeed, as already remarked, this particular case could be dealt with in a very simple manner without making any use of the concept of partial waves. Thus, finally, we are left with the phenomena arising from the incidence of light on the boundary between two material media. Huyghens' construction explains the geometric laws of reflection and refraction in so natural and convincing a fashion that it is difficult to resist the conclusion that his concept of partial waves is well-grounded and is a physical reality in these particular cases.

7. THE SO-CALLED PRINCIPLE OF HUYGHENS

It will be evident from what has been said above that the ideas of Huyghens were not correctly understood or appreciated by later writers. It is not surprising therefore that the whole of the vast literature which was subsequently published and which claims to base itself on the ideas of Huyghens, in reality proceeds on a different basis altogether. This is evident from the fact that the mathematicians whose objective was to develop a " Rigorous

Formulation of the Principle of Huyghens" concerned themselves with precisely the case in which Huyghens' concept of partial waves has no physical meaning or justification, namely the undisturbed propagation of waves from a source situated in a structureless and uniform continuum.

The well-known formula developed by Kirchhoff is an illustration of the foregoing remarks. Here, the disturbance due to the source at the point of observation is expressed as an integral taken over the area of a closed surface within which the point of observation is included but not the source. Each elementary area of the surface appears in the formula as a source from waves diverge with amplitudes which vary with the direction of emission. The line joining the source and the point of observation is also the direction of maximum amplitude for the waves radiated by the element of area which lies on that line *between* them, and of zero amplitude for an element of area which also lies on the same line but on the *opposite side.* Kirchhoff's formula as actually developed refers to the case of sound-waves, and the attempts made to extend it to the case of light have not met with success. But our present concern is not with the mathematics of the formula but with the physics of the subject. The association of the formula with the name of Huyghens—honoured as the founder of the wave-theory of light—has naturally disposed whole generations of physicists to look upon it with favour. It has, however, been made clear by the foregoing remarks that Kirchhoff's approach to the subject is quite different from that of Huyghens. We have, therefore, to ask ourselves: Is Kirchhoff's formula really meaningful? Has it any claim to validity or acceptance considered from the standpoint of optical theory? We shall proceed to consider these questions.

As has already been remarked, one of Huyghens' striking successes is his explanation of the geometric laws of reflection and refraction. His concept of partial waves takes its clearest and most acceptable form in this case, *viz.*, that each element of area of the physical boundary acts as a source of partial waves. Since these move with different velocities in the two media, they should be considered as distinct. In other words, the partial waves in each medium are hemispherical, and it becomes a meaningful physical problem to determine the dependence of the amplitude of the waves with direction on the surface of these hemispheres. It would presumably be a maximum in the direction of the normal to the boundary and zero in directions parallel to the boundary. On the other hand, the very generality of Kirchhoff's formula indicates that it has no physical validity or significance. For, it is not possible to discover or assign any reason why an element of area set at an arbitrary orientation in a continuous structureless medium

should function as a source of secondary waves with specific features related to that orientation. If the concept of partial or secondary waves is at all to be meaningful, the waves should have a physically recognizable origin, *e.g.*, a local discontinuity in physical properties. In its absence, the formula ceases to have any physical content. Kirchhoff's formula thus reveals itself to be a mathematical abstraction which is not relevant or valid in relation to the actual problems of physical optics.

8. SUMMARY

The " Principle of Huyghens " has played an important role in physical optics and especially in the explanation of the phenomena of the diffraction of light. An examination of Huyghen's own ideas as expounded in his original treatise is therefore of interest and is undertaken in the present memoir. It emerges that the particular form given to the " Principle " by mathematical analysts, notably Kirchhoff, does not find any support or warrant in the writings of Huyghens. The concept of " partial waves " introduced by Huyghens in his treatise is based on specific physical considerations. These are totally lacking in relation to the free and uninterrupted propagation of waves in a continuous structureless medium. The latter case is considered in Kirchhoff's theory and it follows that the Kirchhoff formula is a mathematical abstraction which has no relevance or validity in relation to the actual problems of physical optics.

Reprinted from "The Proceedings of the Indian Academy of Sciences"
Vol. XLIX, No. 5, Sec. A, 1959.

THE OPTICS OF MIRAGES

BY C. V. RAMAN AND S. PANCHARATNAM
(*Memoir No.* 115 *from the Raman Research Institute, Bangalore*-6)

Received May 6, 1959

1. INTRODUCTION

CHAPTER IV of the celebrated *Traite de la Lumiere* by Christiaan Huyghens published in the year 1690 deals with " the Refraction of the Air ". Huyghens discussed the effects which arise by reason of the atmosphere having a refractive index varying with the height above the surface of the earth and also with the physical conditions. The problem was considered by him from two distinct points of view, *viz.*, that light consists of rays travelling out from the source, alternatively of wave-fronts propagated through the medium; he showed that the two approaches give completely identical results. Each small piece of the wave-front in an inhomogeneous but isotropic medium moves forward normally to itself with the velocity appropriate to the refractive index at its position; the pieces thus moving forwards join up to form a continuous wave-front. As the result, we find that the rays of light traverse curved paths which are everywhere normal to the wave-fronts.

Mirages constitute one of the most remarkable effects arising from a variation of refractive index of the atmosphere. A very familiar type of mirage is that seen when the light rays are nearly parallel to the heated surface of the earth; the sky and objects on the horizon then appear to be reflected by the surface or rather by the cushion of hot air in contact with it. The phenomenon is frequently observed over asphalted or concrete roads on sunny days and over firm smooth sands in warm weather. Very striking pictures of such mirages are reproduced in a book by Minnaert.[1] Mirages can also be artificially reproduced in the laboratory. R. W. Wood[2] in his *Physical Optics* gives photographs of the phenomenon as thus obtained.

On an examination of the explanations usually given for the formation of mirages, it becomes apparent that they are inadequate and indeed unsatisfactory. The authors felt that the subject should be investigated more fully both from the theoretical and the experimental standpoints. The present memoir records the results of such study and throws new light on the nature and origin of mirages.

2. THE NATURE OF THE PROBLEM

Consider two media of refractive indices μ_1 and μ_2 which we suppose in the first instance to have a sharp plane boundary of separation that is normal to the z-direction. If a plane wave from the first medium be incident on the boundary at a glancing angle ϕ_1 then, in general, there would exist both a reflected and a refracted wave; the former follows the usual law of reflection, while the direction ϕ_2 of the latter is given by Snell's law of refraction $\mu_1 \cos\phi_1 = \mu_2 \cos\phi_2$, the intensities of the reflected and refracted beams being given by Fresnel's formulæ. If however $\mu_1 \cos\phi_1 > \mu_2$, there exists no refracted beam, so that the incident energy is totally reflected.

In the present paper, we are interested in the case in which the change from the refractive index μ_1 to the refractive index μ_2 does not occur abruptly at a sharp boundary, but takes place continuously and progressively over a certain transition layer—the refractive index being constant along planes normal to the z-direction. In this case, general considerations based on wave-optics show that the refracted wave in the second medium should still satisfy Snell's law of refraction $\mu_1 \cos\phi_1 = \mu_2 \cos\phi_2$ while, if the refractive index does not change appreciably within distances of the order of the wavelength, there would normally be no reflected wave in the first medium, the entire energy going into the refracted wave. However, an exception to to this statement must clearly arise when $\mu_1 \cos\phi_1 > \mu_2$. Since in these circumstances, there can be no refracted wave in the second medium, a reflected wave must be formed, the reflection being in fact total. Such a situation occurs in the phenomenon of the mirage. Though it is clear that the reflected wave in the first medium must be a plane wave following the usual laws of reflection, the nature of the disturbance within the transiton layer can only be ascertained by a rigorous analytical treatment from wave-theoretical considerations. This has been done by Epstein[3] choosing a particular profile for the refractive index variation and the solution obtained is expressed in terms of hypergeometric functions; however, the only application of Epstein's work which he has made and is relevant to the present problem is to remark that there is total reflection in the conditions mentioned, the angle of incidence being equal to the angle of reflection.

3. SOME ELEMENTARY CONSIDERATIONS

The theory of the mirage which is usually accepted purports to base itself on geometrical optics. Any incident ray initially making an angle ϕ_1 with the plane of the stratifications would be progressively deviated from its course so that when it reaches a limiting layer of refractive index

$\mu_l = \mu_1 \cos \phi_1$, it would be tangential to that layer. It is then usually stated that the ray continues to curve round, but it must be noted that on the basis of geometrical optics the ray should really continue parallel to the stratifications along the limiting layer. This would happen to every one of the rays in the incident beam. As a consequence, the entire energy of the incident radiation would be concentrated in an infinitely thin layer having the refractive index μ_l. If we take two adjacent rays a finite distance apart, the part of the wave-front between them would on entry into the transition layer swing round and at the same time contract in its extension and ultimately become reduced to a point travelling along the limiting layer of refractive index μ_l.

It is evident that the foregoing situation deduced from elementary considerations presents some fundamental difficulties. In the first place, as we have already noted, considerations of a general nature indicate that the energy reaching the limiting layer which cannot penetrate further must of necessity be returned back through the transitional region into the first medium as a reflected wave. Then, again, as has been stressed by Planck[4] in his *Theory of Light*, the equivalence of the results of geometrical optics with the exact results of wave theory is subject to certain limitations. One of them is that the amplitude of the wave-function must be only slowly variable in space. This condition is clearly violated in the present case, for according to what has been said above, the disturbance would be infinitely great at the limiting layer of refractive index μ_l and zero immediately below it. It follows that while the actual facts may bear some resemblance in their general features to this elementary picture, the real situation would nevertheless be rather different; firstly, a reflected wave must emerge from the limiting layer; secondly, there should exist a finite intensity below the limiting layer, while above that layer, the intensity though large would exhibit successive maxima and minima arising from the interference of the incident and reflected waves and progressively diminish as we move from the limiting layer toward the first medium.

It is clear from the foregoing remarks that the situation in the vicinity of the limiting layer would resemble that well-known in physical optics to appear in the vicinity of caustic surfaces formed by reflection or refraction of light. The effects observed in such cases are usually described as due to the propagation of a cusped wave-front along the caustic surface. Such a result is clearly capable of experimental test in the present case and has been confirmed by the observations made by us presently to be described.

4. Observation of the Caustic and accompanying Interferences

The arrangement first tried for producing the mirage under laboratory conditions was similar to that described by R. W. Wood. A long steel plate, approximately $1\frac{1}{2}$ metres long, 10 cm. in width and 2 cm. in thickness was supported horizontally and heated from below by four gas burners, the upper surface being covered with soot. The arrangement did not prove very satisfactory since the rising of the hot air from the heated surface prevented a sufficiently sharp temperature gradient from being established, while at the same time such phenomena as could be observed were of an exceedingly labile nature owing to the turbulent convection of air near the hot surface. However, the conditions were greatly improved when an electric fan was used to blow away the hot air from above the surface. When this was done, it was possible to see clearly the reflected image of the 'mountain peaks' formed by holding the serrated edge of a cardboard against an illumined ground glass screen at the far end of the plate. The action of the fan in sharpening the temperature gradient was picturesquely revealed by suitably lowering the eye, when a luminous thin cushion of air was seen to form over the heated surface. In order to study the phenomena critically, the object viewed was replaced by an illuminated slit kept parallel to the heated surface and the light diverging from it was rendered parallel by a collimating lens and allowed to fall obliquely on the hot plate. The beam was allowed to cover the whole length of the plate, the angle of incidence being adjusted merely by moving the slit. Further, the plate was heated electrically to ensure a fairly uniform temperature. When the temperature was sufficiently high and the glancing angle sufficiently small, the field at the nearer end of the plate, when examined through a pocket lens, displayed a bright strip of light separated by a dark region from the plane of the hot surface. It was clear, however, that the use of the fan was causing a smearing out of the caustic and preventing phenomena of the nature of interferences from being clearly seen.

The above difficulty was removed by the artifice of turning the plate edgewise so that the hot surface was now vertical, though its length remained horizontal. The hot air now flowed up in streamlines parallel to the surface of the plate, thus rendering the use of the fan unnecessary. With the slit now vertical, it was possible to observe clearly with the aid of a pocket lens a bright caustic bordered by a large number of interferences of the type referred to in the last section. A photograph of this phenomenon is given in Plate XII, Fig. 1. In order to completely arrest the oscillations of the caustic and its fringes, a short exposure (1/100th of a second) was necessary so that the slit had to be brilliantly illuminated. Accordingly, sunlight was

used for this purpose and this, in fact, enabled also the introduction of a red filter for increasing the clarity of the fringes.

It will be seen that the field in Plate XII, Fig. 1 consists of three parts. To the right of the bright caustic (*i.e.*, towards the heated surface) the field is dark, while to its left lies an illumined strip containing a large number of interference fringes whose separation narrows down to a constant value as we move away from the caustic; this is of course due to the increasing inclination between the two branches of the cusped wave-front. Since the heated plate is necessarily of finite extension, the reflected part of the cusped wave-front does not extend to infinity but is terminated. This manifests itself in the field of view by the occurrence of a second edge to the left of which the intensity is considerably less (though not zero), the edge being bordered by some broad fringes.

5. Relation of the Mirage to the Cusped Wave-Front

The actual mirage is observed when the eye is kept at any point which lies on the bright strip of light lying to the left of the caustic, the eye being focussed on the plane containing the object, *i.e.*, at infinity. It is to be expected that two images would then be seen whose positions lie respectively along the directions of the normals drawn from the eye to the two branches of the cusped wave-front leaving the nearer edge of the plate. In fact, the fringes observed to the left of the caustic in Plate XII, Fig. 1 may be regarded as due to the interference between the light from two such virtual sources, the progressive narrowing of the spacing of the fringes to the left of the caustic corresponding to the increasing separation of the sources. That the separation of the two images observed depends on the position of the eye or aperture through which the phenomenon is viewed is illustrated in Plate XII, Fig. 3. In order to make the nature of the image evident, the serrated edge of a hacksaw blade has been used to form one of the edges of the slit. An aperture was kept in front of the lens and the succession of photographs exhibit the alteration in the phenomena as the aperture is gradually moved to the left. When the aperture is on the bright caustic, a single image is seen, while as it is moved to the left this separates into two images, one of which is a direct (or more properly, a refracted) image, the second being an inverted reflected image, the separation between the two gradually increasing. A remarkable feature of the sequence of phenomena illustrated in Plate XII, Fig. 3 is the occurrence of a third erect image close to the reflected image in the fourth and fifth photographs of the sequence; this image starts developing when the aperture has been moved towards the outer edge of the central illumined strip (where broad fringes start appearing

in Plate XII, Fig. 1), becoming coincident with the reflected image when the aperture is exactly at the edge mentioned. As the aperture is moved further left, only the 'direct' image continues to be visible, as is shown in the last photograph of the sequence; this is to be expected since the reflected part of the cusped wave-front is no longer received through the aperture.

The third image mentioned above can be cut off by inserting an opaque screen near the farther end of the heated plate and adjusting it so that its edge protrudes a little beyond the surface of the plate. This shows that the image is due to the ordinary refraction of rays directly entering the edge of the heated layer at the farther end of the plate. The main features of the path of such rays may be deduced from the experimental observations described in the previous paragraph. The terminus of the reflected part of the cusped wave-front corresponds to certain limiting rays entering the region at the farther end of the plate. Rays which are able to enter the edge of the heated layer at a closer distance to the plate than these limiting rays proceed a longer distance before emerging from the heated stratum and also suffer a larger deviation. These rays give rise to the erect third image; in fact, the second refracted wave-front, obtained by drawing the surfaces orthogonal to these rays, meets the termination of the reflected part of the wave-front so as to form a second cusp.

It may be mentioned that an erect third image has been observed and photographed by Hiller[5] in his study of the mirage produced by a long vertical wall. He has attributed the phenomenon to multiple reflection, but it is more probably due to ordinary refraction as in the present case.

Till now we have dealt mainly with the case when the distant object is of negligible angular dimensions. When an object of finite angular dimensions is used, the point on the image which corresponds to any particular point on the object is to be determined as before for each setting of the eye. In this case there will be a distortion of the images because the position of the limiting layer as well as the inclinations of the cusped wave-fronts corresponding to any particular point on the object, varies with the position of the object-point. Plate XII, Fig. 2 shows the photographs taken using as the object a small model of a bird made of glass. This was placed near the focal point of the collimating lens. The sequence of photographs show the variation in the appearance of the phenomena as the eye is moved away from the plate. The appearance of a third erect image in addition to the usual reflected image may be discerned in the last two photographs of the sequence.

6. Mathematical Treatment

Let the z-axis be taken normal to the direction of the stratifications of refractive index, the plane of incidence being taken as the xz plane. The disturbance ψe^{iwt} at every point must satifsy the wave equation

$$\nabla^2\psi + k^2\mu^2\psi = 0$$

Since the refractive index is a function only of z, we may put $\psi = e^{ikpx}u(z)$ where u satisfies the one-dimensional wave equation

$$\frac{d^2u}{dz^2} + Q^2(z)\,u = 0 \tag{1}$$

where

$$Q(z) = k\sqrt{\mu^2 - p^2} \tag{2}$$

We shall take the layer at which the refractive index becomes equal to p to be the plane $z = 0$, and the law of variation of refractive index to be given by

$$\left.\begin{array}{lll} \mu^2 = \mu_1^2 & \text{for} & z \geq z_1 \\ \mu^2 = a^2z + p^2 & \text{for} & z \leq z_1 \end{array}\right\} \tag{3}$$

The general solution of the wave equation (1) for the region $z \geq z_1$, can be written

$$u = A\,\{\exp. - ik\,(\sqrt{\mu_1^2 - p^2}z + \theta) + R\exp. ik\,(\sqrt{\mu_1^2 - p^2}z + \theta)\} \tag{4}$$

If the incident wave be given by $A\exp. - ik\,(\mu_1\sin\phi_1 . z + \theta)$, the first term will represent the incident wave with $p = \mu_1\cos\phi_1$ so that from (3) the plane $z = 0$ is the limiting layer; the second term represents a reflected wave with the angle of reflection equal to the angle of incidence, R being the coefficient of reflection. We shall take R to be a real quantity, since this can be ensured by appropriately choosing the phase factor θ, the complex reflection coefficient being $Re^{2i\theta}$.

Writing $ka = 3\gamma/2$ the wave equation for the region $z \leq z_1$ can be written

$$\frac{d^2u}{dz^2} + \left(\frac{3}{2}\gamma\right)^2 zu = 0 \tag{5}$$

This is one of the transformations of Bessel's equation, the general solution being expressed in terms of the cylinder functions of order one-third:

$$u = z^{1/2}\,C_{1/3}\,(\gamma z^{3/2})$$

i.e., with the usual notation for the Bessel functions

$$u = \xi^{1/3} \{A_1 J_{1/3}(\xi) + B_1 J_{-1/3}(\xi)\} \tag{6}$$

where $\xi = \gamma z^{3/2}$.

For negative values of z the above may be rewritten to avoid the use of imaginary arguments, by putting

$$\xi = i \mid \xi \mid = i\gamma \mid z \mid^{3/2}$$

$$u = \xi^{1/3} \{A_1 I_{-1/3}(\mid \xi \mid) - B_1 I_{1/3}(\mid \xi \mid)\}$$

$$= \xi^{1/3} \left\{A_1 \frac{2}{\pi} \sin \frac{\pi}{3} K_{1/3}(\mid \xi \mid) + (A_1 - B_1) I_{1/3}(\mid \xi \mid)\right\}$$

Since $I_{1/3}(\mid \xi \mid)$ tends to infinity as $z \rightarrow -\infty$, the condition that u remain finite below the limiting layer gives $A_1 = B_1$. Hence below the limiting layer $z = 0$, we have

$$u = \frac{\sqrt{3}}{\pi} \mid \xi \mid^{1/3} A_1 K_{1/3}(\mid \xi \mid) \tag{7}$$

while above the limiting layer

$$u = A_1 \xi^{1/3} \{J_{1/3}(\xi) + J_{-1/3}(\xi)\} \tag{8}$$

For layers not close to the limiting layer, the asymptotic expansion of (8) for large z may be used:

$$u \sim \left(\frac{3}{2\pi}\right)^{1/2} \xi^{-1/6} A_1 \cdot 2 \cos\left(\xi - \frac{\pi}{4}\right)$$

$$\sim \left(\frac{3}{2\pi}\right)^{1/2} \xi^{-1/6} A_1 \left\{\exp. - i\left(\xi - \frac{\pi}{4}\right) + \exp. i\left(\xi - \frac{\pi}{4}\right)\right\} \tag{9}$$

The constants A_1 and θ may be determined by the condition that the solutions (4) and (8) must join smoothly, *i.e.*, by the condition that u and du/dz must be continuous at $z = z_1$. Since the solution u as given by (8) is a real quantity below the plane z_1, it is clear from (4) that the reflection coefficient R must be unity, as is to be expected. The constants A_1 and θ may be determined with sufficient accuracy by equating separately the amplitudes and arguments of the periodic functions in (4) and (9) since it is the periodic functions which contribute mainly to du/dz.

Thus

$$A_1 \sim \left(\frac{2\pi}{3}\right)^{1/2} \xi_1^{1/6} A$$

$$\theta \sim \left(\xi_1 - \frac{\pi}{4}\right) - \mu_1 \sin \phi_1 z_1 \tag{10}$$

where $\xi_1 = \gamma z_1^{3/2}$.

Before proceeding to discuss the solution obtained, we may remark that the function ξ in terms of which the solution is expressed can be given a physical interpretation. Using the fact that $\xi = \int_0^z Q dz$, it may be easily verified that

$$[\xi(z) + kpy] = \int_0^P \mu ds$$

where the right-hand side represents the optical distance from the point $P(x, z)$ to the limiting layer, measured along the limiting ray of geometrical optics. Thus in (9) the solution has been represented as the sum of two disturbances, the surface of constant phase of one of the wave-fronts being taken to be that given by geometrical optics; the other reflected wave-front is symmetrical to the first but has suffered a phase change of $\frac{1}{2}\pi$. The asymptotic solution (9) does not apply near the limiting layer as may also be seen by the fact that it gives an infinite intensity at that layer. The correct solution (8) applicable near the limiting layer may be written in a form similar to (9)

$$u = A_1 \xi^{1/3} \frac{J_{1/3}(\xi) + J_{-1/3}(\xi)}{2\cos(\xi - \frac{1}{4}\pi)} \{\exp. - i(\xi - \tfrac{1}{4}\pi) + \exp. i(\xi - \tfrac{1}{4}\pi)\} \tag{11}$$

Thus, the foregoing mathematical treatment provides a justification for the qualitative description given earlier in the paper of the optical field as a cusped wave which moves along the limiting layer.

Using equations (7) and (8) and a table of Bessel functions, the function $(u/A_1)^2$ which is proportional to the intensity has been plotted against $|\xi|^{2/3}$ which is proportional to the distance z from the limiting layer. The graph is shown in Fig. 1 in the text. It is seen that slightly above the limiting layer there is a large concentration of intensity. Above this again we have a series of interferences which progressively diminish in spacing and intensity, while below the maximum, the intensity rapidly falls to zero.

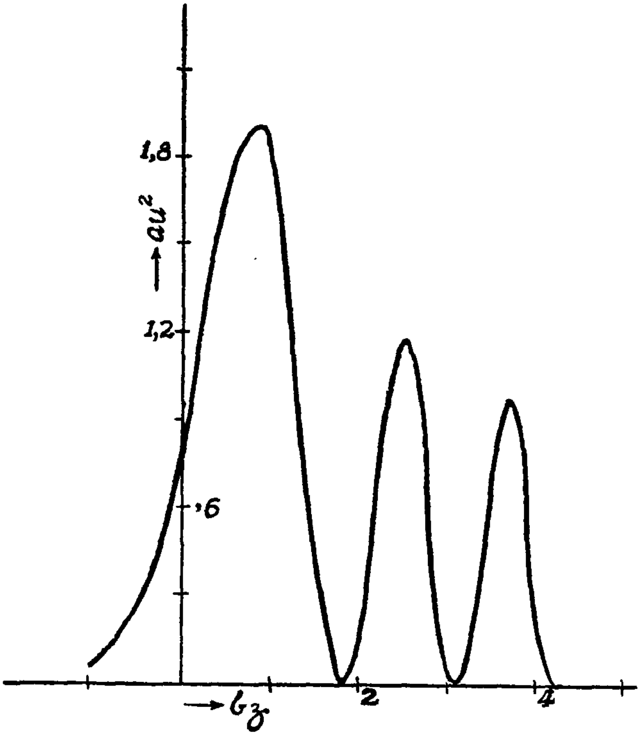


FIG. 1

It will be seen that the method of solution adopted is similar to that adopted in quantum mechanics where similar mathematical problems arise. In fact, though we have for simplicity considered only the case when μ^2 varies linearly, a similar method[6] can be adopted even in the general case, the solution being again expressible as the sum of two Bessel functions if we neglect the square of the wavelength.

7. SUMMARY

Experimental studies reveal that when a collimated pencil of light is incident obliquely on a heated plate in contact with air, the field of observation exhibits a dark region adjacent to the plate into which the incident radiation does not penetrate, followed by a layer in which there is an intense concentration of light and then again by a series of dark and bright bands of progressively diminishing intensity. Photographs of these features have been obtained and are reproduced in the paper. The observed facts find a satisfactory explanation when the problem is considered from the standpoint of wave-theory and more completely on the basis of a formal analytical treatment. The optical characters of the mirage are observed to stand in the closest relation to the wave-optical phenomena referred to, changing with

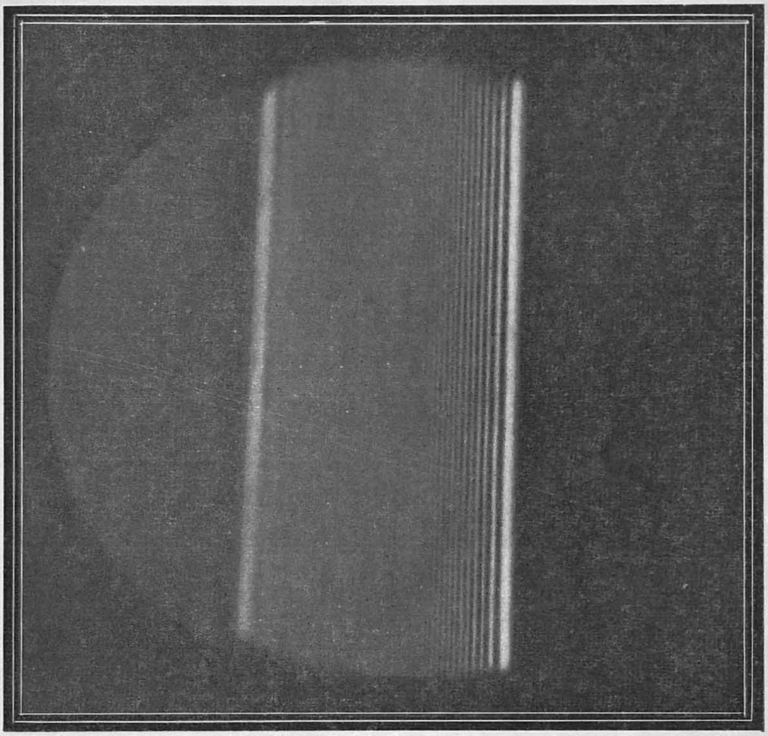


FIG. 1

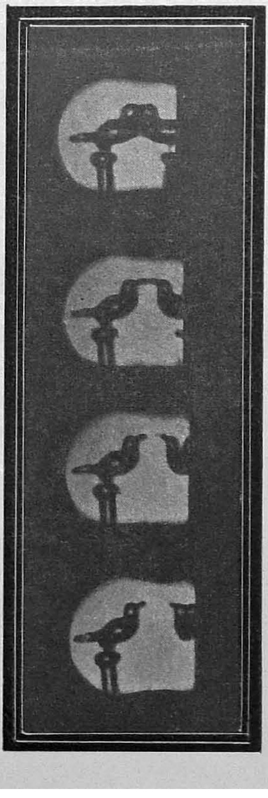


FIG. 2

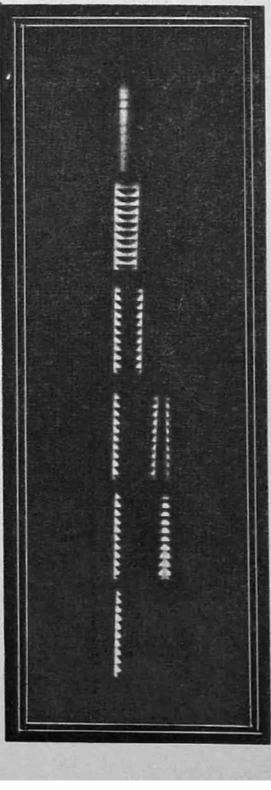


FIG. 3

the position of the observer's eye in the field of interference. The explanations of the mirage usually given are thus seen to be inadequate and unsatisfactory.

8. References

1. Minnaert, M. .. *Light and Colour in the Open Air*, G. Bell and Sons, London, 1940, Plate V.
2. Wood, R. W. .. *Physical Optics*, The Macmillan Co., New York, 1911, Plate III.
3. Epstein, S. .. *Proc. Nat. Acad. Sci.*, U.S.A., 1930, **16**, 628.
4. Planck, M. .. *Theory of Light*, Macmillan and Co., London, 1932, 204.
5. Hiller, W. .. *Phys. Zs.*, 1913, **14**, 718; 1914, **15**, 303.
6. Morse and Feshbach .. *Methods of Theoretical Physics*, McGraw-Hill, 1953, Part II, 1096.

Reprinted from "The Proceedings of the Indian Academy of Sciences",
Vol. XLIX, No. 6, Sec. A, 1959

CAUSTICS FORMED BY DIFFRACTION AND THE GEOMETRIC THEORY OF DIFFRACTION PATTERNS

BY SIR C. V. RAMAN
(*Memoir No.* 116 *from the Raman Research Institute, Bangalore*-6)

Received May 27, 1959

1. INTRODUCTION

IT has long been known that the diffraction patterns of the *Fraunhofer class* exhibit, in certain simple cases, a recognizable relation to the form of the apertures giving rise to them. For example, the pattern of a triangular aperture resembles a six-rayed star; that of an elliptic aperture exhibits elliptic rings of the same eccentricity, but turned through a right angle, and so on. Investigations[1, 2] by the present author and his collaborators made and published many years ago revealed that geometric relationships which are essentially of the same nature are also exhibited by diffraction patterns of the *Fresnel class.* It was shown in the publications quoted that in the Fresnel patterns of an elliptic aperture there appears a concentration of the diffracted light along the evolute of the cross-section of the emergent pencil. The effect is in the nature of a caustic and indeed exhibits the characters usually associated with caustics. A similar effect is observable in the Fresnel patterns of elliptic discs, and indeed quite generally in the diffraction patterns of apertures or obstacles with undulating boundaries. The origin of the effect was further elucidated by investigations[3, 4, 5] made by the author and his co-workers in which the technique of the Foucault test was employed to determine the characters of the radiation from diffracting boundaries in various circumstances. It was observed that the elements of a diffracting boundary act as sources of secondary radiations which appear with maximum intensity in a plane normal to themselves; from each element, *two streams* of diffracted radiation emerge which are in opposite phases and appear respectively on either side of the direction of incidence of the light on the boundary.

From the facts stated, it follows that it should be possible to develop a theory of diffraction patterns on a purely geometric basis. Indeed, this was done successfully at the time[6] for the case of the Fraunhofer pattern of a semicircular aperture—the so-called heliometer diffraction figures. That it could also be done for *Fresnel patterns* was evident, but the matter was not

followed up until many years later. The basic ideas of the theory were expounded in the second chapter of the author's "Lectures on Optics" written out and printed in the year 1943 but not then given out for publication. Basing himself on the ideas set out therein, Y. V. Kathavate[7] worked out in detail the cases of a great many Fresnel diffraction patterns and his results were published in a series of six papers which appeared together in the April 1945 issue of these *Proceedings.* Many excellent photographs were reproduced with his papers which exhibit a highly satisfactory concordance with the results of the geometric theory. Elliptic obstacles and apertures give Fresnel patterns of great beauty and afford perhaps the most striking examples of the power and elegance of the geometric method in treating diffraction problems. Kathavate had indeed dealt with these cases in his fifth paper, but it was felt that this work had not covered them adequately and that it would, therefore, be useful to consider them afresh. That indeed is the object of the present paper.

2. OBSERVATION OF THE PATTERNS

The simple device of holding a circular aperture or disc obliquely in the path of a beam of light enables us to obtain the equivalent of an elliptic aperture or screen of the desired eccentricity. It is necessary, for obtaining the best results, that the metallic sheet carrying the aperture or the metallic disc should be very thin and that the circular edge has been accurately machined and has a smooth polish. The patterns depend very much on the eccentricity; by progressively altering the tilt, the entire sequence ranging from a circular aperture or disc to an elliptic aperture or disc of high eccentricity can be readily studied. Kathavate's paper reproduces a series of photographs obtained in this manner of the Fresnel patterns observed in the region of shadow of discs and in the region of the light transmitted by apertures. His photographs, however, do not show the phenomena outside the region of the transmitted light in the case of elliptic apertures of moderate or large eccentricity. To fill this gap, the series of six photographs reproduced in Plate XIV were recorded in the present investigation. An aperture of 5 mm. diameter in a very thin steel sheet was used. The source of light was a pinhole illuminated by sunlight filtered through a red glass. The source was at a distance of 136 centimetres from the aperture and the divergent beam emerging therefrom was received on a photographic plate held 168 centimetres away from it. The tilt of the aperture was progressively altered, being quite large for Fig. (*a*) and quite small for Fig. (*f*). The exposures given were very short in the case of Figs. (*e*) and (*f*). They were much larger for Figs. (*b*), (*c*) and (*d*), while quite a long exposure was given for Fig. (*a*) recorded with the largest tilt.

Special interest attaches to the study of the transition from the Fresnel to the Fraunhofer class of diffraction pattern in the case of elliptic apertures. Accordingly, the series of eight photographs reproduced in Plate XV was obtained to exhibit this transition. The patterns were recorded with the same aperture of 5 mm. diameter set at the same obliquity in all the cases. The source of light was a fine pinhole illuminated by sunlight filtered through a red glass. A long-focus telescope objective of about 5 feet focal length was used to form a focussed image of the pinhole some 30 feet away from the objective on the other side. The light passed through the objective and then through the aperture and was received on a photographic plate held at an appropriate distance behind the aperture. This distance in the series of eight photographs was 18 inches, 3, 6, 15, 19, 22, 26 and 30 feet respectively.

For the convenience of the reader, two of the plates of Kathavate's paper are reproduced with the present communication as Plates XVI and XVII respectively. These, together with Plate XIV and XV recorded in the present study, form the experimental material which the reader of the paper can compare with the results of the theory.

3. Geometric Theory of the Patterns

The geometric theory of diffraction patterns was expounded in the second chapter of the author's "Lectures on Optics" already referred to. The theory was further discussed in these *Proceedings* by G. N. Ramachandran[8] and also by Kathavate in the first of his series of six papers. Its application to numerous cases by graphical methods was also worked out by Kathavate. It is therefore unnecessary to go over the same ground here. For our present purpose, we commence by considering the simple case of a Fresnel pattern of a circular disc as it illustrates the ideas on which the theory is based very clearly. As is well known, the shadow thrown by a circular disc held in a divergent pencil of light exhibits a central spot surrounded by concentric dark and bright rings. The entire pattern can be deduced if we recognize that the edge of the disc is itself the source of the diffracted radiations which enter the region of the shadow. The radiations from the entire circumference arrive at the centre in the same phase and accordingly give the maximum possible intensity at that point. As we move away from the centre of the pattern in any direction, the elements of the boundary at the two ends of a diameter of the disc become effectively the sources of the radiation reaching the point of observation. The distances traversed by them being different, a phase difference arises and accordingly we would observe interferences. As has been shown by G. N. Rama-

chandran,[8] the positions of the circular rings around the central bright spot can be quantitatively deduced from these considerations.

With appropriate modifications, arguments of the same general nature enable us to deduce the characters of the diffraction pattern in the shadow of an elliptic disc. This pattern arises from the radiations diffracted by the boundary and entering into the region under consideration. It is useful to have a drawing in which the elliptic boundary of the shadow is shown and a series of normals are drawn to it. We proceed on the basis that each element of the diffracting boundary of the actual screen sends out radiations having the maximum intensity in the plane normal to itself and hence may indeed be regarded as effective only in that plane. Hence, to find the part of the boundary which is operative at any given point of observation on the screen, we draw a normal from that point to the boundary of the shadow. The length of the normal thus drawn is indicative both of the amplitude and the phase of the radiation reaching the point of observation from the corresponding point on the actual boundary. The amplitude decreases rapidly as the length of the normal increases, while the phase alters progressively. By superposing the radiations from the various points on the boundary which are operative at any particular point of observation, we derive the effects observable at that point. Thus finally, the features of the pattern over the entire area can be envisaged.

Two diagrams have been drawn in the manner indicated above. Figure 1 in the text below refers to the case in which the ratio of the major axis to the minor axis of the ellipse is 4: 3. The evolute of the ellipse in this case falls entirely within the ellipse itself. Figure 2 in the text below refers to the case in which the ratio of the major to the minor axis is as 2: 1. In the latter case, only part of the evolute and its cusps situated on the major axis appear within the ellipse, while the rest of the evolute along with the two other cusps lies outside the ellipse on a prolongation of the minor axis. The following features emerge from these diagrams. From any point outside the evolute, only two normals can be drawn to the elliptic boundary. On the other hand, from points within the figure of the evolute four normals can be drawn to the boundary. At any point on the evolute itself, these become coincident in pairs and hence here again, effectively only two normals can be drawn.

Thus from the diagrams we can draw the following general inferences. Within the figure of the evolute we have four sources operative situated at the points on the boundary, the superposition of which will give us a characteristic interference pattern. At points outside the evolute only two sources

on the boundary would be operative and these again would give us an interference pattern of a relatively simpler type. The diagrams show, however, that special situations arise at and near the four cusps of the evolute and also

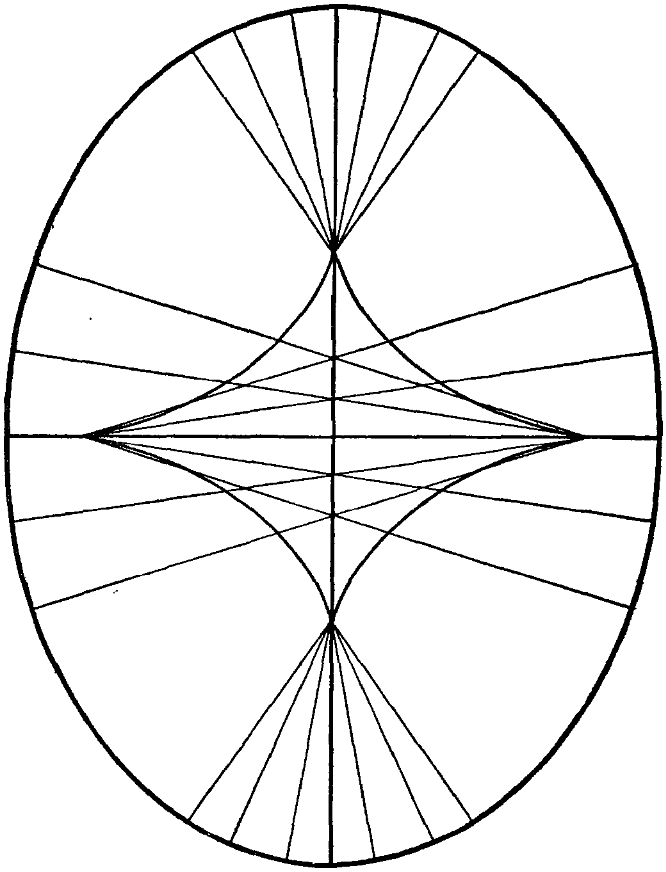


FIG. 1

along the evolute itself. The two cusps on the *major axis*, as can be seen from the diagram, are effectively the foci of the diffracted radiations from a considerable length of the boundary on either side of the terminations of the major axis. It follows that the resultant intensities at these two cusps must be very large. A similar situation would also appear at the two cusps on the *minor axis*. But since the convergence of the diffracted rays is less and the angles of diffraction are larger in the latter case, the intensity of illumination at the cusps on the minor axis would be less pronounced and would also diminish rapidly with increasing ellipticity of the boundary. The evolute being itself the locus of the foci of the diffracted radiations, it would be a region in which the resultant intensity is large. Further, since the radiations reaching their foci at the cusps would diverge again, there would be a noticeable concentration of light on the prolongation of the minor axis beyond the cusps located on it.

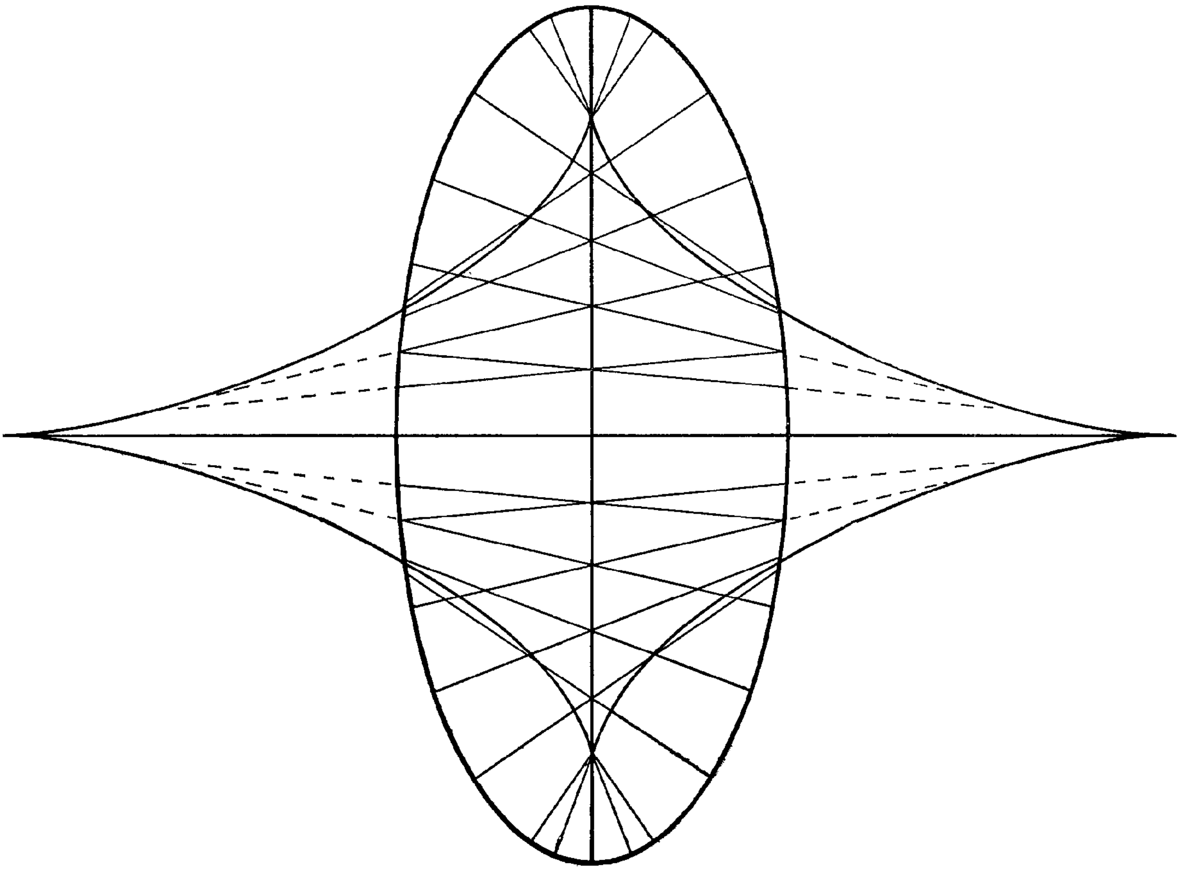


FIG. 2

4. THE FRESNEL PATTERNS OF ELLIPTIC DISCS

We now proceed to consider the nature of the interference patterns observable within the shadow of an elliptic disc. As has already been remarked, these would be of an altogether different nature in the region within the evolute and outside it. Provided the ellipticity of the disc is not too great, the situation in the region outside the evolute would be analogous to that in the shadow of a circular disc; in other words, we would have interference rings which are closed curves appearing in that region. But they would be neither circular nor of uniform intensity. It is easily seen that an interference ring of any given order would cut the minor axis at points further away from the centre of the pattern than its intersections with the major axis. The interference rings would therefore be of approximately elliptic form, their major axis coinciding with the minor axis of the shadow and *vice versa.* They would be most conspicuous at and near the points where they cross the minor axis and scarcely visible where they cut across the major

axis, the reason for such difference being that the radiations which interfere would be approximately of equal intensity in the former case and very unequal in the latter case.

We now proceed to consider the interference pattern inside the region of the evolute. At and near the centre of the pattern, four sources of diffracted radiation would be effective which lie on the boundary at or near the ends of the major and minor axes of the ellipse. The radiations from the ends of each axis would arrive at the centre in the same phase, but such phase would be different for the two axes and would also vary as we move away from the centre in either direction. Hence the interference pattern due to the superposition of the effects of the four sources would be an array of spots arranged in regular order along equidistant lines parallel to the major axis of the ellipse, and similarly also on lines parallel to the minor axis. Since the ends of the major axis are further away from each other and from the point of observation than the ends of the minor axis, the array of spots on lines parallel to the major axis would be spaced wider apart and would also be more conspicuous than the arrays on lines parallel to the minor axis. It should also be remarked that since two normals can be drawn to the boundary from any point on the evolute itself and two sources of diffracted radiation are therefore effective thereon, the evolute itself would not appear as a continuous curve of high luminosity but would be cut across by the interference bands parallel to the major axis referred to above.

A few remarks will suffice regarding the case in which the elliptic disc is of large eccentricity. A substantial part of the evolute would then lie outside the shadow and we are not concerned with it here. It is obvious that in these circumstances, the diffraction pattern within the shadow would exhibit prominently a set of interference bands which are for the most part equidistant and parallel to the major axis but which would be splayed out near their ends, as the sources of diffracted radiation giving the interferences then approach each other. The pattern would be crossed by another set of bands parallel to the minor axis, but these would be much closer together and far less conspicuous than the bands parallel to the major axis, since the sources of diffracted radiation giving rise to them are further apart and effectively much weaker in the region under consideration.

The reader may compare the foregoing results with the patterns reproduced in Plate XVI which exhibit the patterns within the shadow of elliptic discs of six different ellipticities.

5. THE FRESNEL PATTERNS OF ELLIPTIC APERTURES

Two distinct regions have to be considered in the discussion of Fresnel patterns of elliptic apertures. These may be designated as the internal and external regions respectively. In the first or internal region the radiations diffracted by the boundary are superposed on the geometric pencil of light emerging through the aperture and the pattern observed is determined by their interferences. The second or external region is the region of shadow which contains only the diffracted radiations from the boundary. The discussions given earlier for the case of an elliptic disc are helpful in dealing with the present problem. Indeed, the results of that discussion may be taken over *mutatis-mutandis* in considering the internal pattern; we have to remember, however, that the phases of the diffracted radiations would all be reversed and that in the final result the interferences with the transmitted pencil of light have also to be taken into account.

The diffraction pattern in the exterior region may be deduced in a simple fashion. If the ellipticity is small or moderate, the pattern would consist of closed interference curves. These would approximate to ellipses in their configuration, their major axis being parallel to the minor axis of the aperture and *vice versa.* Kathavate verified this result experimentally (see Plate XI of his paper). It is evident for reasons already explained that the visibility of the rings would be much greater in the vicinity of the minor axis than of the major axis.

If the ellipticity be considerable, the figure of the evolute and its prolongation beyond the two cusps on the minor axis would extend into the region of the shadow. This extension would be then the most noticeable feature of the external diffraction pattern, since the intensity would be greatest along the figure of the evolute. This is evident in Figs. (*a*), (*b*), (*c*) and (*d*) in Plate XIV. These also exhibit other features which we shall proceed to consider. We may remark in the first place that a few interference fringes are to be noticed alongside the evolute running parallel to it. This is a general feature of all caustics and does not therefore need special explanation. In the region enclosed by the evolute and also cutting across it at regular intervals, a regular sequence of bands appears, these being transverse to the minor axis of the ellipse and exhibiting a marked curvature. They arise from the superposition of the diffracted radiations from the two sides of the ellipse and are conspicuous because the interfering radiations are of comparable intensity. The narrowing of the ellipse as we move from its centre towards the ends causes the curvature of the bands. They are most intense at small angles of diffraction and fade away as we proceed towards the cusps of the

evolute. Finally, it may be remarked that outside the figure of the evolute and cutting across its prolongation on the minor axis beyond the cusps, the elliptic rings referred to above should also manifest themselves, though only very feebly.

We have now to consider the nature of the internal patterns in which the radiations diffracted by the boundary and the incident pencil of light as it emerges from the aperture are superposed. It is evident that the radiations diffracted through small angles by the elements of the boundary and interfering with the primary beam would give rise to fringes forming closed curves adjacent to the boundary and running parallel to it. Further inside the patterns, however, the disposition of the interferences would be profoundly altered, firstly by the convergence of the diffracted radiation to foci at the four cusps of the evolute, and elsewhere by the superposition of the diffracted radiations from two or even four elements of the boundary at particular points within the patterns. Since at the two cusps on the major axis the intensity of the diffracted radiation is large and indeed comparable with that of the primary beam, the interferences with the latter at and near these foci should be very conspicuous; a disposition of the fringes in their vicinity along arcs of circles centred around them is also to be expected. Then again, the superposition of the radiations from opposite sides of the elliptic boundary would result in cancelling out the influence of the curvature of that boundary on the disposition of the interferences within the pattern. Such an effect should be particularly conspicuous in respect of the parts of the boundary which run nearly parallel to the major axis of the ellipse. Thus, we should expect the fringes of the usual type running parallel to the elliptic boundary to be replaced by a series of straight bands parallel to the major axis, their number and disposition being determined by the ellipticity and other relevant circumstances. Since the effect of four sources is superposed on the primary beam at and near the centre of the pattern, the bands parallel to the major axis would in that region be crossed by another set of bands parallel to the minor axis which would be less conspicuous and also more closely spaced.

The foregoing results are very clearly illustrated by the patterns reproduced as Figs. (*e*) and (*f*) in Plate XIV and by the array of fifteen pictures in Plate XVII taken from Kathavate's paper. The latter were recorded by him with elliptic apertures of different sizes and eccentricities. They were obtained by making three sharp circular holes in a metal sheet, with radii respectively equal to 6·9 mm., 5·3 mm. and 4·8 mm. By tilting the sheet with respect to the incident beam, three diffraction patterns could be photographed simultaneously. Thus with five different tilts, fifteen photographs

in all were obtained. In the reproduction, all the fifteen figures were enlarged to nearly the same size.

6. Transition from the Fresnel to the Fraunhofer Patterns

The Fresnel pattern of an elliptic aperture observed in its immediate vicinity is an elliptic patch of light orientated in the same manner as the aperture itself. The Fraunhofer pattern observed at the focus of the lens which concentrates the light passing through the aperture is also an elliptic patch of light, but turned round through a right angle and surrounded by elliptic rings of the same shape and orientation. The successive stages of this transformation are illustrated by the series of eight photographs reproduced in Plate XV. The features seen in the photographs are readily understood in terms of the preceding discussion of the nature of the Fresnel patterns themselves. As we proceed from the aperture towards the focus of the lens, the cross-section of the geometric pencil diminishes progressively, and what has been referred to as the interior pattern contracts to a point and ultimately vanishes. What survives is the exterior pattern and it is the progressive alteration in the nature of that pattern that is represented in the series of photographs. In Fig. 2 of Plate XV we see the evolute and its accompanying fringes outside the elliptic patch of light. An important role is played by the very faint extension of this evolute beyond its cusps. A brush of light extends outwards along the minor axis; as can be seen from the series of photographs, this is interrupted by interferences crossing it, due to their intersection with the elliptic rings which are a characteristic feature of the exterior patterns; these are however much too feeble to give any other evidence of their presence in the first four photographs of the sequence. They are barely visible in the fifth photograph, but are much better seen in the sixth. The seventh and the eighth photographs show the final transformation of the brush into a complete set of elliptic rings.

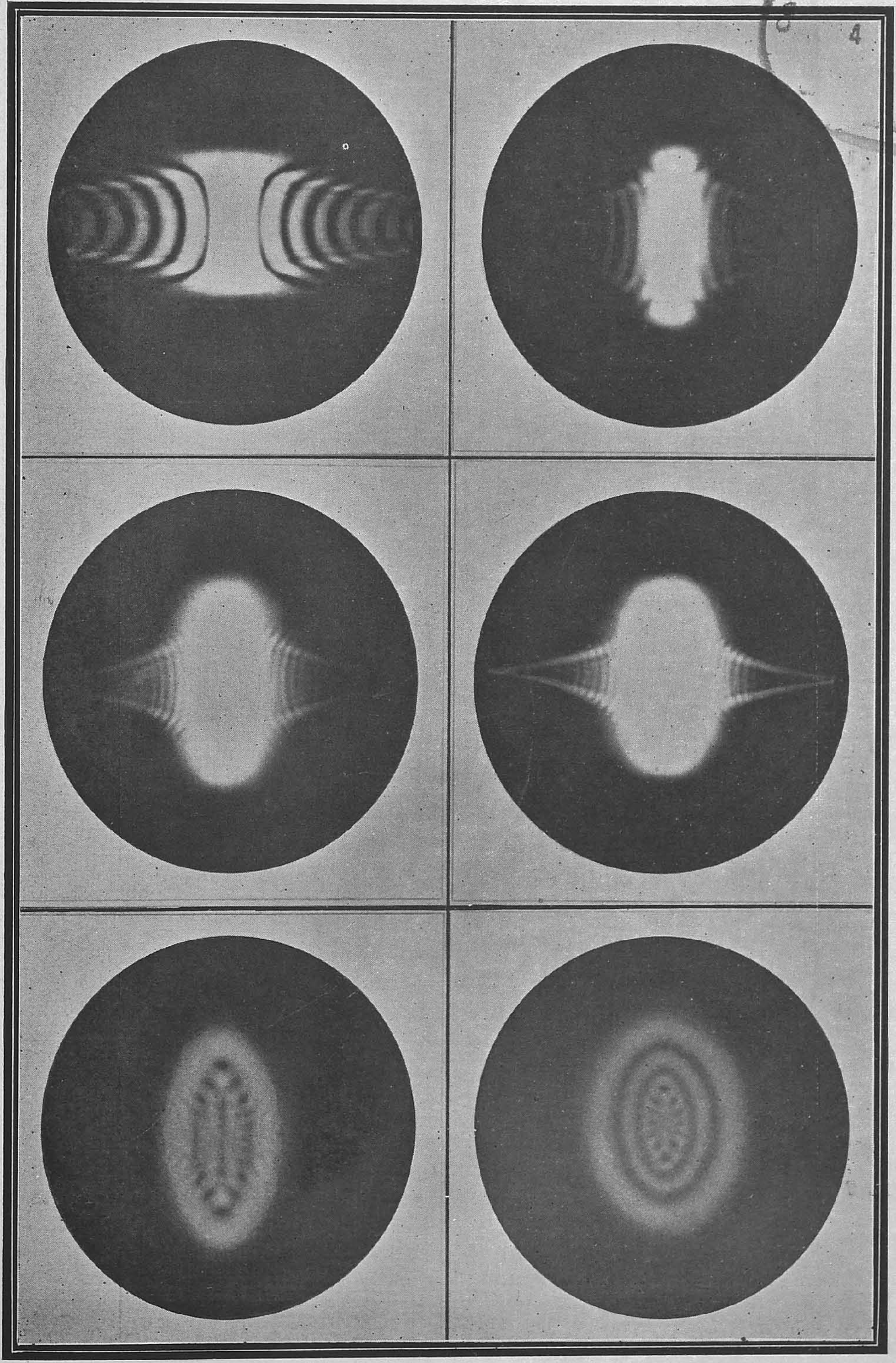
7. Summary

The Fresnel patterns of elliptic discs and apertures afford a striking example of the power of the geometric method in treating the theory of diffraction patterns. They are discussed in detail and a series of photographs exhibiting the features indicated by the theory are reproduced.

8. References

1. Raman, C. V. .. *Phys. Rev.*, 1919, **13**, 259.
2. Mitra, S. K. .. *Phil. Mag.*, 1919, **38**, 289.
3. Raman, C. V. and Ghosh, P. N. *Nature*, 1918, **102**, 205.

PUBLIC 4 JAN196



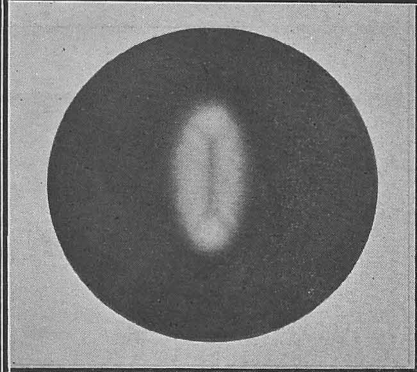
(*a*) (*b*)

(*c*) (*d*)

(*e*) (*f*)

Fresnel Diffraction Patterns of Elliptic Apertures.

FIG. 1



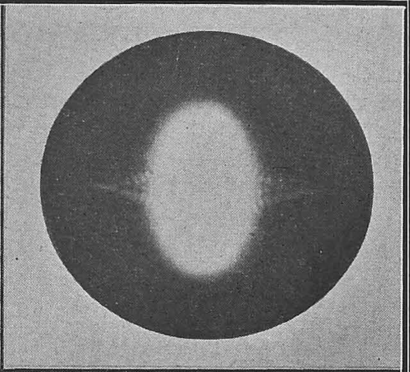
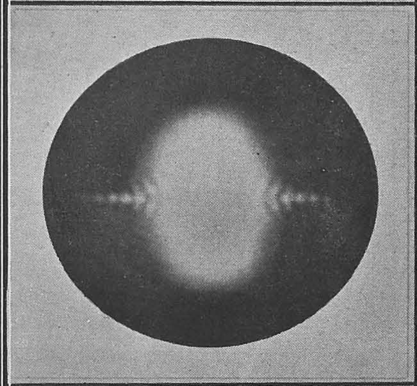


FIG. 2

FIG. 3



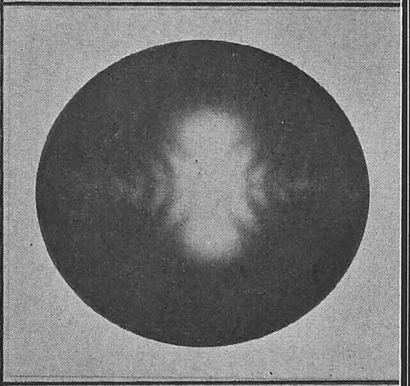
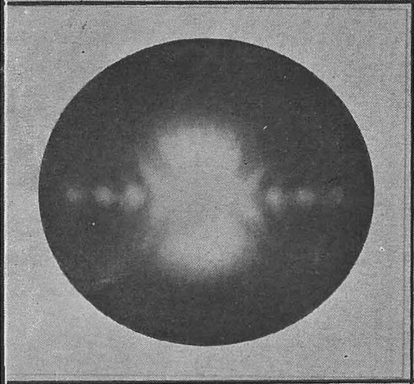


FIG. 4

FIG. 5



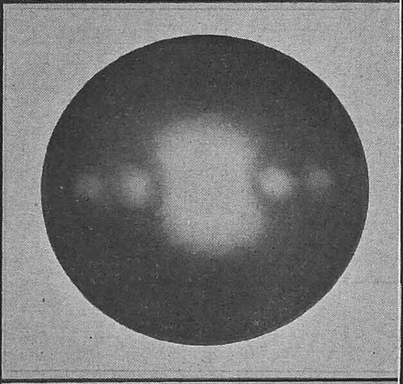
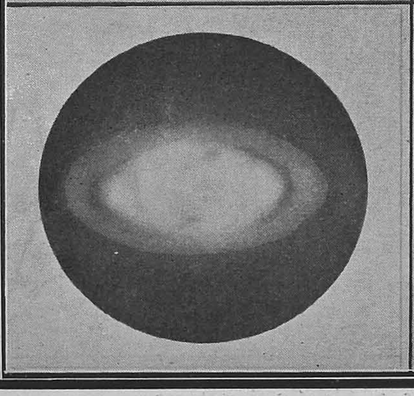


FIG. 6

FIG. 7



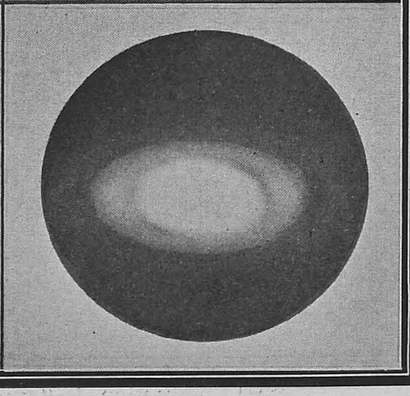
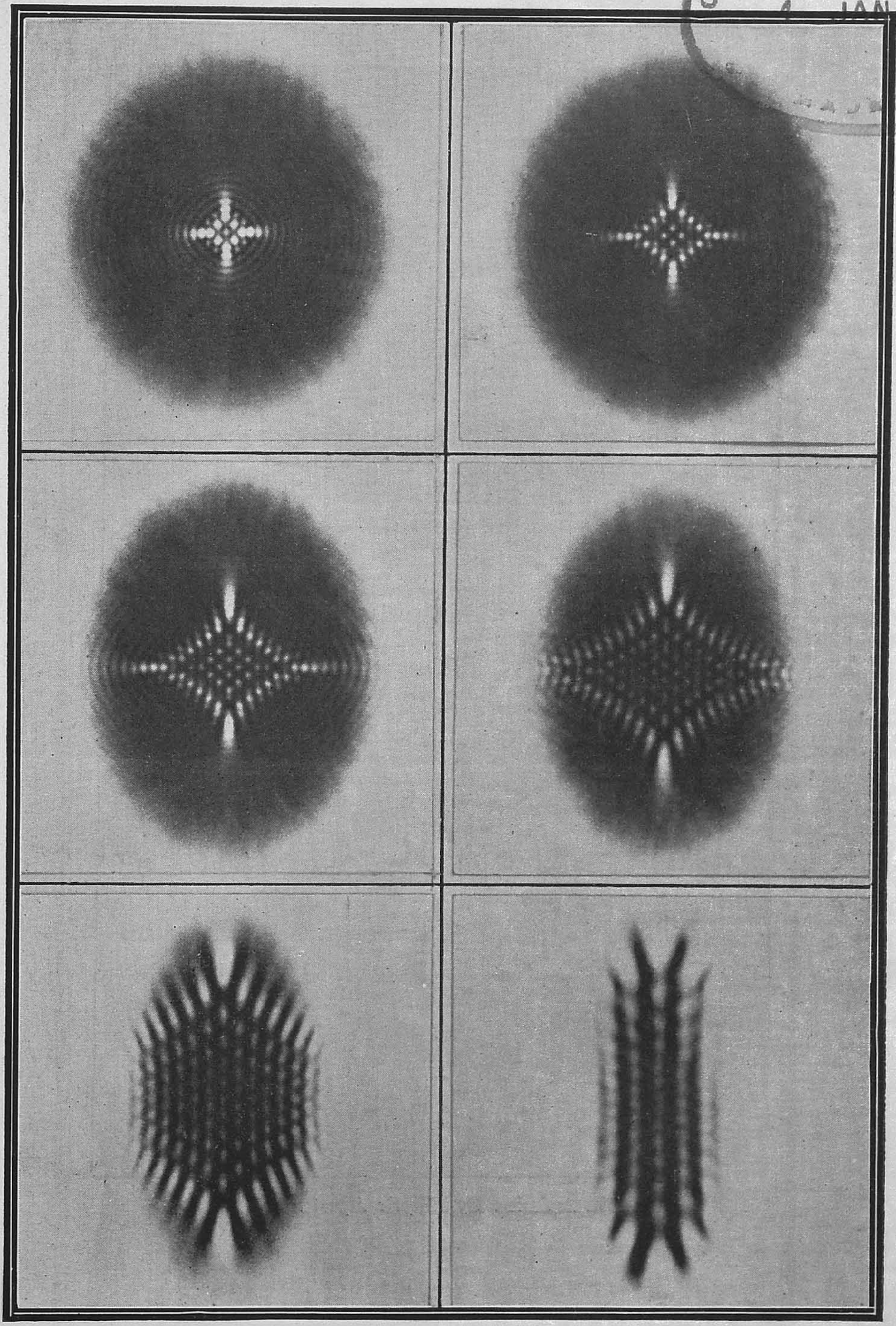


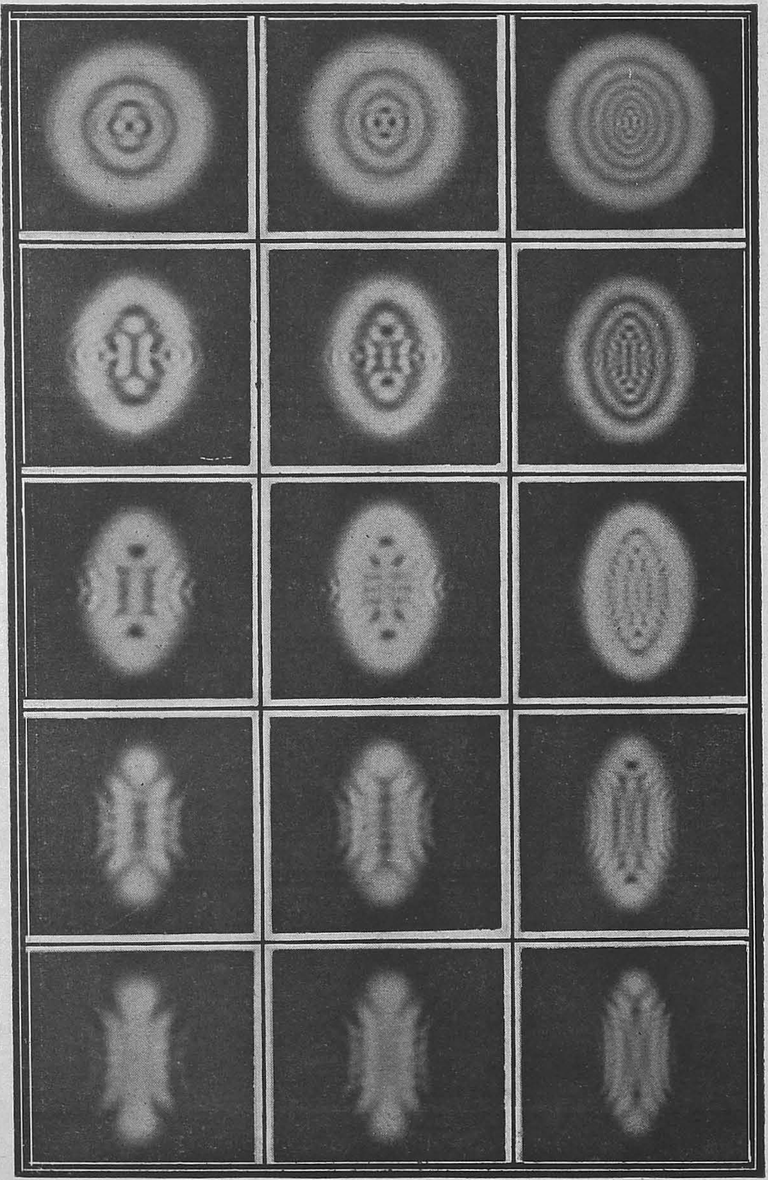
FIG. 8

Transition from Fresnel to Fraunhofer Patterns.

CONNEMARA PUBLIC LIBRARY
4 JAN 1965
MADRAS



Pattern in the Shadow of Elliptic Discs.



Transmission by Elliptic Apertures.

4. Ghosh, P. N. .. *Proc. Roy. Soc.*, 1919, **96 A**, 257.
5. Banerji, S. K. .. *Phil. Mag.*, 1919, **37**, 112.
6. Mitra, S. K. .. *Proc. Ind. Assocn. Cult. Sci.*, 1920, **6**, 1.
7. Kathavate, Y. V. .. *Proc. Ind. Acad. Sci.*, 1945, **21 A**, 177.
8. Ramachandran, G. N. .. *Ibid.*, 1945, **21**, 165.

Reprinted from "The Proceedings of the Indian Academy of Sciences", Vol. L, No. 2, Sec. A, 1959

HUYGHENS' PRINCIPLE AND THE DIFFRACTION OF LIGHT

Part I. Theoretical Considerations

BY SIR C. V. RAMAN

(*Memoir No.* 118 *from the Raman Research Institute, Bangalore-6*)

Received July 23, 1959

1. INTRODUCTION

THE first three chapters of Huyghens' Treatise on Light published in the year 1690 carry the titles "On Rays Propagated in Straight Lines"; "On Reflection"; "On Refraction". They seek to explain the familiar facts of geometrical optics on the basis of the idea that light consists of waves propagated outwards from the original source in the luminiferous medium. Huyghens based himself on certain hypotheses regarding the nature of the medium and on the nature of the light waves themselves. The adequacy of his explanations of the rectilinear propagation of light and of the geometric laws of reflection and refraction of light has necessarily to be viewed against the background of those hypotheses. The treatise of Huyghens exerted a great influence on the development of the wave-theory of light in later years. Surprisingly enough, however, his ideas were not correctly understood by those who followed him. They ascribed to Huyghens various statements not to be found in his treatise and then proceeded to criticise his ideas in an unjustifiable manner. In a recent memoir by the present author published in these *Proceedings*, a precis was given of the first three chapters of Huyghens' treatise and it was shown that, if we accept the hypotheses of Huyghens regarding the nature of the luminiferous medium and the nature of light, the rest of his argument forms a coherent and complete explanation of the facts which he sought to elucidate.

Treatises on physical optics devote a good deal of space to an exposition of what they refer to as "Huyghens' Principle" and the more advanced treatises proceed to sketch a mathematical development which is described as "The Rigorous Formulation of Huyghens' Principle" and associate it with the name of Kirchhoff. The readers of those treatises are invited to believe that Huyghens' Principle as thus formulated is an adequate basis for

the explanation of the phenomena of the diffraction of light. However, in the author's memoir mentioned above, it was indicated that the explanations referred to are vitiated by the misunderstanding of the ideas of Huyghens on which they claim to be based. It is the object of the present memoir to go more fully into these matters and to show that the concepts of Huyghens can indeed be used as a framework for an elementary theory of diffraction, but that they lead to results which are wholly different from those indicated by the well-known formulæ of Kirchhoff. Extensive experimental studies of the phenomena of the diffraction of light have also been carried out by the author and the results of the same are reported in the second part of this memoir. They vindicate the present theoretical approach and contradict the consequences of Kirchhoff's theory, showing thereby that the latter is unsustainable.

2. THE WAVE-OPTICS OF HUYGHENS

When we examine the first three chapters of Huyghens' treatise, it becomes clear that a certain basic idea regarding the propagation of light underlies them all. What Huyghens sought to show in his first chapter is that in a homogeneous medium, each little piece of the primary wave emerging from a source of light is capable of travelling in a direction normal to itself more or less independently and that the primary wave-front is the locus or surface at which all the little pieces of which it is made up arrive at the same instant. The explanation of the laws of reflection and of refraction given in the second and third chapters proceeds on the same basis. Each piece of the original wave-front from the light-source on reaching the boundary between two media of which the optical properties are different finds itself unable to continue on its original course by reason of that fact. Such a situation arises, for example, if both media admit of the propagation of light through them but with different velocities. In these circumstances, the individual pieces of the wave take fresh paths which are different in the two media; the directions of travel in each case are such that the pieces of the original wave-front which are diverted from their path on reaching the boundary join up together again and form new wave-fronts in each medium. A simple geometric construction enabled Huyghens to ascertain the directions of travel which enable these conditions to be satisfied in each medium, from which again the geometric laws of reflection and refraction follow immediately. Huyghens' derivation of these laws is both simple and convincing. Regarded as a physical theory, it is also highly successful since it demonstrates that the refractive indices of the two media are in the inverse ratio of the velocities of light in them.

The concept of " partial waves " introduced by Huyghens plays a most important role in his arguments and indeed forms the hard core of his theory. It is not introduced *ad hoc* but is put forward with a definite physical justification wherever it is brought in and made use of. In the first chapter of the treatise in which the rectilinear propagation of light is sought to be explained, the partial waves considered are those which arise by reason of the structure which the luminiferous medium is assumed to possess. Each individual particle of the medium when disturbed by the passage of the primary wave becomes a source from which partial waves spread out in all directions, but these partial waves by reason of their excessive feebleness can produce a sensible effect only when a great number of them arrive simultaneously at a given point of observation: this again is only possible when the sources of the partial waves all lie on the straight line joining the original source and the point of observation. Accordingly, it is justifiable to regard the primary wave as having travelled out along that line and identify it with the summation of the partial waves of which the effects are superposed at the point of observation.

The partial waves considered in the second and third chapters of the treatise have a different origin. Here, Huyghens found himself compelled to introduce the idea that the elementary *areas* of the surface of separation between two media having different optical properties become sources of partial waves: the waves which go back into the first medium build up the reflected wave-front, while those which go forward into the second medium build the refracted wave-front.

3. The Nature of Diffraction Phenomena

When we speak of the diffraction of light, we have in mind certain effects which are observed when the free propagation of light is modified or influenced by the presence of obstacles in its path. It is clear that the nature of the obstacles, including their optical properties and their configuration in space, would determine these effects and it follows that the factors referred to would play the leading role in any theory of diffraction. If, bearing this in mind, we seek to discover in the ideas of Huyghens a possible approach towards an understanding of the phenomena of diffraction, it becomes apparent that no such approach can be found in his explanation of the rectilinear propagation of light. On the other hand, his theory of reflection and refraction does offer itself as a basis. For, as already stated, it makes use of the idea that each element of area of the boundary between two media on which light is incident is a source of partial or secondary waves in the two media. Conceptually, these waves can diverge from each element

in various directions, but the requirement imposed by the theory of Huyghens that the disturbances originating at the different elements of area should arrive simultaneously at a common wave-front fixes the actual direction of their movement. If, instead of considering light waves as simple pulses, we take account of their periodic nature and of the possibility of interferences between the secondary or partial waves having their origin at the different elements of area on the boundary between the two media, the restriction of the observable effect to precisely defined directions ceases to exist. In other words, the diffraction of light becomes a possibility.

Thus, we arrive at the important conclusion that a theory of diffraction which makes use of Huyghens' concept of partial waves has to base itself on the waves of that nature which arise in association with the reflection and refraction of light at the boundary between two media with different optical properties.

4. THE CHARACTER OF THE SECONDARY WAVES

Accordingly, we proceed to consider the partial or secondary waves having their origin at the elements of area of a boundary between two media of different refractive indices on which light is incident. There would clearly be two sets of such secondary waves travelling out respectively into the two media. The velocity of travel and the amplitude of the disturbance in the two sets being different, they must be considered as completely distinct from each other. If both media are isotropic, the configuration of the secondary waves in each medium would be hemispheres. The particular circumstances of the case, *viz.*, the refractive indices of the two media, the angle of incidence of the primary waves on the boundary and the state of polarisation of the incident light would determine the manner in which the energy of the incident radiation would be divided up between the reflected and refracted wave trains. The same circumstances would also determine the amplitude of the disturbance in the secondary waves sent out respectively into the two media.

The theory of diffraction proceeds by considering the secondary radiations emitted in different directions by the elements of area of the reflecting or refracting boundary and summing up their effects at the point of observation, having regard to their amplitudes and phases at that point. The first step in the theory is to write down a formal expression for the partial or secondary waves. Their amplitude would clearly be proportional to the amplitude of the primary disturbance incident on the boundary and reflected or refracted by it, as the case may be, and hence we have to find a dimensionless magnitude which expresses the proportionality. The quantities which might appear in it include, firstly, the element of area dS on the boundary,

r the radius vector joining the element with the point of observation, ϕ the angle between r and the normal to the plane of the boundary, and λ the wave-length of the light. It is reasonable to assume that the solid angle subtended by the element of area dS at the point of observation would determine its effect at that point. This solid angle is $dS \cdot \cos \phi / r^2$. If we multiply this by the quantity r/λ, we obtain the dimensionless number $dS \cos \phi / \lambda r$, which exhibits an inverse proportionality both to r and to λ; the inverse proportionality with respect to r is to be expected for waves diverging outwards hemispherically, while the inverse proportionality with respect to λ is an indication that we are dealing with a wave-optical effect.

The important result emerging from the above is the appearance of $\cos \phi$ in the expression for the amplitude of the secondary waves. We shall refer to it as the *obliquity factor.* It has the value unity in the direction of the normal to the reflecting or refracting boundary, while it vanishes in any direction which lies in the plane of that boundary. It is evident that if we are concerned with a reflecting or refracting boundary of finite area and if the point of observation is at a sufficiently great distance from it, the angle ϕ may, without sensible error, be assumed to be the same for all the elements of area. It is evident also that in such a case, when we proceed to investigate the nature of the diffraction pattern at such distant point by summing up the effects of the elements of area with due regard to their phases and squaring the resultant amplitude to obtain the observed intensity, the square of the cosine of the obliquity, *viz.*, $\cos^2 \phi$ would appear in it as a multiplying factor.

The result stated above is obviously of very general validity in respect of the diffraction patterns of the Fraunhofer class observed in various circumstances. All that is required is that the diffraction arises by reason of the limitation of the area of a plane surface at which light is reflected or through which it is transmitted; in the case of reflection, the material may be either a dielectric or a metal. It is not necessary that the surface should be continuous or that it should have uniform reflecting or transmitting power over the entire area. It might, for example, consist of several parallel strips, thus forming a plane diffraction grating. Further, since refraction at the boundary between two media which differ only infinitesimally in refractive index is equivalent to a simple transmission, it follows that the result would also be applicable to diffraction patterns of the Fraunhofer class arising from the passage of light through apertures in thin opaque screens.

5. Verification of the Obliquity Law

Any elementary treatment of diffraction theory can only be expected to be valid when the linear dimensions of the diffracting aperture are large

compared with the wave-length of the light. As the angular spread of the diffraction pattern would in these circumstances be small, an experimental test of the law of the secondary wave might seem impracticable. Fortunately, however, this is not the case. For, the angle of diffraction ϕ is measured from the direction of the normal to the aperture and hence when the incidence of the light on the aperture is oblique, ϕ may be large enough for the factor $\cos^2 \phi$ to vary rapidly over the area of the diffraction pattern. Further, at such settings the diffraction patterns are spread out over a fairly wide angular range even when the dimensions of the aperture are many times larger than the wave-length. In these circumstances, the effect of the $\cos^2 \phi$ factor on the distribution of the intensity in the pattern becomes conspicuous and can indeed easily be observed and measured.

We may illustrate these remarks by considering a simple case, *viz.*, a diffracting aperture which is a plane strip bounded by parallel straight edges. As is well known, when the effects due to the infinitesimal elements of such an aperture are summed up, the expression obtained for the intensity in its Fraunhofer pattern includes a factor of the form $\sin^2 \zeta/\zeta^2$. This factor has a maximum value when $\zeta = 0$, and vanishes when $\zeta = \pm \pi, \pm 2\pi, \pm 3\pi$, etc. Since the value of $\sin^2 \zeta/\zeta^2$ is unaltered by a reversal of the sign of ζ, the graph of the function when set out with ζ as the abscissa is a symmetric curve in which the maxima on either side intermediate between the zero values are of equal intensity. The factor $\cos^2 \phi$ by which the expression for the intensity is multiplied would, however, modify this situation to an extent determined by the circumstances of the case.

In the particular case of normal incidence of the light on the aperture, $\zeta = \pi a \sin \phi/\lambda$, a being the width of the aperture, λ the wave-length and ϕ the angle of diffraction as already defined. More generally, when the light is incident on the aperture at an angle θ in a plane normal to its edges, $\zeta = \pi a (\sin \phi - \sin \theta)/\lambda$. Differentiating this, we obtain $d\zeta = \pi a/\lambda . \cos \phi \, d\phi$. Hence, as the incidence is made more oblique and $\cos \phi$ diminishes in value, the angular spread of the pattern determined by the increments of $d\phi$ becomes greater. The bands for which ϕ is greater than θ would also appear more widely spaced than those for which ϕ is less than θ. In these circumstances, the multiplying factor $\cos^2 \phi$ would have a very conspicuous influence on the distribution of intensity in the pattern. The bands for which ϕ is greater than θ would be much less intense than those for which ϕ is less than θ; indeed as ϕ approaches the limiting value $\pi/2$, the intensity in the former cases would become vanishingly small.

6. The Results of Experimental Study

The present theory of diffraction and that of Kirchhoff thus differ fundamentally in the observable results which they indicate. This is scarcely a matter for surprise since they approach the diffraction problem from completely different points of view. Whereas the diffracting body plays the leading role in the present theory, it is not considered at all in the Kirchhoff formulation; the latter is based on the idea that the primary radiation from a source situated in free space can be expressed as a summation of secondary radiations from a set of sources distributed over a closed surface in free space enclosing the point of observation. The present theory leads to the result that the amplitude of the secondary waves vanishes along the plane of the surface at which they originate and increases progressively as we move away from that plane towards the direction of its normal. On the other hand, the Kirchhoff formulation indicates that the secondary waves have a maximum amplitude in the forward direction of the light rays from the original source and zero amplitude in the backward direction. The difference between the consequences of the two theories is of such a striking character that it is a simple matter by means of experimental study to decide between them.

In view of the importance of the issue here raised for a correct understanding of the theory of the diffraction of light, numerous experimental studies have been carried out by the writer. Diffracting apertures of various sizes ranging from several centimetres down to fractions of a millimetre have been employed. The angles of incidence of the light on the apertures have been varied from normal incidence up to grazing incidence. The circumstances in which the diffraction manifests itself have also been varied to include various cases, *e.g.*, the reflection of light at a plane surface of a dielectric or metal, the emergence of light after refraction through a transparent medium at various angles, the internal reflection of light within a transparent medium at various incidences and the transmission of light through apertures in thin opaque screens. The cases investigated include both simple and multiple apertures and plane diffraction gratings prepared by various techniques and operating by reflection as also by transmission. In all the cases investigated, the consequences of the present theoretical approach are completely vindicated by the facts of observation.

7. Summary

The conventional treatment of diffraction problems based on the so-called Principle of Huyghens as analytically formulated by Kirchhoff is based on a misunderstanding of the original ideas of Huyghens regarding the propagation of light. It seeks to express the luminous effect due to the primary

source as a summation of the effects of secondary sources situated on the elementary areas of a surface in free space enclosing the point of observation. Since the diffraction of light is a consequence of the presence of obstacles in the path of the light waves, the optical character of the obstacles and their configuration in space are of the very essence of the problem. It follows that the approach adopted in the Kirchhoff theory is misconceived and erroneous. It is however possible to base a theory of diffraction on Huyghens' concept of partial waves and his explanation of the reflection and refraction of light at the boundary between media with different optical properties. But this leads to results which differ fundamentally from those indicated by Kirchhoff's formula. It is shown that the issue which thus presents itself is readily capable of experimental test.

Reprinted from "Curr. Sci.", August 1959, 28, 309-313

THE OPTICS OF MIRAGES

Sir C. V. Raman

(Memoir No. 118 of the Raman Research Institute, Bangalore-6)

1. Introduction

In describing and interpreting our visual impressions, it is natural to adopt the geometric approach provided by ray-optics, since this is the most familiar and therefore the most easily understood approach. But this does not mean and it would indeed be incorrect to assume that even the most familiar optical effects which we can perceive with the unaided vision can always be explained on the basis of the ideas of geometrical optics. The propagation of light in various circumstances needs in general the concepts of the wave-theory for its correct and complete elucidation. Only in the particular case of a set of waves of constant type and the rays associated with them are the wave-optical and ray-optical descriptions completely equivalent.

The foregoing remarks have a bearing on the phenomena of atmospheric optics, in other words of the effects noticed when distant sources of light are viewed through great depths of air. The atmosphere of the earth is an inhomogeneous medium and though the refractivity and its variations are both small, this is set off by the very long optical paths often involved. Hence, the deformations of the wave-fronts in their passage from the light-source to the observer are by no means negligible and in particular circumstances, their consequences may be very conspicuous. The familiar phenomenon of the twinkling of the stars may be cited as an example. In the case of terrestrial objects, we are concerned with the lower levels of the atmosphere where the refractivity is a maximum and also liable to large variations by reason of a variety of circumstances. Hence, a variety of effects may arise which are readily observable. As remarked above, it would not be correct to assume that these effects can all be understood or explained on the basis of the ideas of geometrical optics. It is, on the other hand, to be anticipated that the aid of the wave-theory would be needed for the purpose.

2. The Origin of Mirages

Mirages constitute a remarkable group of effects arising from variations of the refractive index of the atmosphere. They are generally described as of two kinds, being respectively the so-called superior and inferior types of mirage. The latter are quite common and may be described as the manifestation above the level heated surface of the earth of a reflection of the sky and other elevated objects: the latter appear as inverted images against the background of the reflected sky and thus simulate the reflection of terrestrial objects at the surface of a pool or lake. The superior type of mirage arises when the thermal conditions of the atmosphere are the reverse of those which give rise to mirages of the inferior type; they are observed when the atmosphere rests on a cold level surface above which there lies a hot stratum of air. Objects at or near the level of the cold surface are usually visible to the observer, and in addition inverted images of them are also seen higher up in the atmosphere. Pictures of both types of mirage are to be found reproduced in many treatises on optics and meteorology. A famous example of the superior type of mirage in which objects thus seen reflected were ships at sea was that described and figured by Vince in the year 1799. His drawings are reproduced or referred to in many accounts of the subject. We shall have occasion to refer to them again later in this article.

3. Mirages and Ray-Optics

The theory of mirages usually accepted purports to base itself on geometrical optics. This explanation is said to have been first put forward by Monge, but it was later elaborated and discussed by many other authors. The interested reader will find a very full review of the literature covering 80 pages in the Second Edition (1922) of Pernter's treatise on meteorological optics as revised by Exner. The explanations usually given are illustrated by

drawings which show the curved path of the rays from the source which reach the observer and are perceived by him as a reflected image of the source. These drawings seem very plausible, but when one examines them with care, it becomes evident that the explanations put forward do not really come to grips with the problem and may indeed be described as a kind of make-believe. That a geometric theory is fundamentally incapable of explaining the phenomenon of the mirage becomes clear on a critical examination of the subject.

The circumstances in which mirages are observed bear a superficial resemblance to those in which the familiar phenomenon of total reflection occurs, *viz.*, light travelling in a medium of higher refractive index meets a medium of lower index at an incidence exceeding the critical angle and is then turned back. There is, however, a fundamental difference between the problem of the mirage and the circumstances of total reflection referred to above. We do not have in the atmosphere anything in the nature of a discontinuous change of refractive index; what we are actually concerned with is a progressive change of index. In the latter circumstances, a pencil of rays travelling obliquely through the stratified medium would, according to Snell's law of refraction, be progressively deviated until it reaches a layer at which its course becomes tangential to the plane of the stratifications; thereafter, it would continue on a course parallel to the stratifications. No question of total reflection can therefore arise.

4. Mirages and Wave-Optics

If it be assumed that the wave-optical and ray-optical descriptions of the behaviour of light are completely equivalent, it would follow that the rays and wave-fronts in an isotropic but inhomogeneous medium are everywhere normal to each other. Accordingly, if the course of the rays is known, the wave-fronts form a set of surfaces cutting them orthogonally. It was stated above as a consequence of Snell's law of refraction that a ray of light initially making an angle with the plane of the stratifications would be progressively deviated from its course until it becomes tangential to that plane and would then continue on a course parallel to the stratifications. This would happen to every one of the rays of an incident beam of light. Accordingly, if we take two adjacent rays at a finite distance apart, the part of the wave-front between them would swing round and at the same time contract in its extension and ultimately become a point which moves parallel to the stratifications in a plane whose position can be specified exactly for the particular circumstances of the case.

An approach to the problem of the mirage based on the ideas of geometrical optics thus leads us to conclude that the energy of the incident radiation would be concentrated at a limiting plane which it reaches but is unable to penetrate. The intensity of illumination in that plane would therefore be infinite. Since such a result is physically inadmissible, it follows that the approach which leads to it should be given up in favour of a different and more rigorous treatment based on the first principles of the wave-theory.

5. The Analytical Theory and Its Results

To make the problem tractable, it has to be idealised to some extent. We consider the optical behaviour of a slab of the medium which is assumed to be of finite thickness and bounded by plane parallel faces extending to infinity. The material is also assumed to be stratified in planes parallel to its faces, the refractive index being μ_1 at the front surface and μ_2 at the rear, μ_1 being greater than μ_2. Plane waves of light are assumed to be incident on the slab at a glancing angle ϕ_1. It is clear that if $\mu_1 \cos \phi_1$ is greater than μ_2, the incident waves cannot emerge from the rear face of the slab. We are interested in ascertaining the nature of the disturbance both within the slab and in front of it in the circumstances stated.

The solution of the problem is contained in a paper by S. Pancharatnam and the present author published in the *Proceedings of the Indian Academy of Sciences* for May 1959. We must be content here with a brief statement of the results. The paper quoted also contains a full account of the results of experimental studies on the subject. This will be drawn upon very freely in the latter half of the present article. It should be mentioned that the experiments, besides confirming the results of the analysis have also revealed other unsuspected features which assist in the elucidation

CONNEMARA PUBLIC LIBRARY
4 JAN 1965
MADRAS

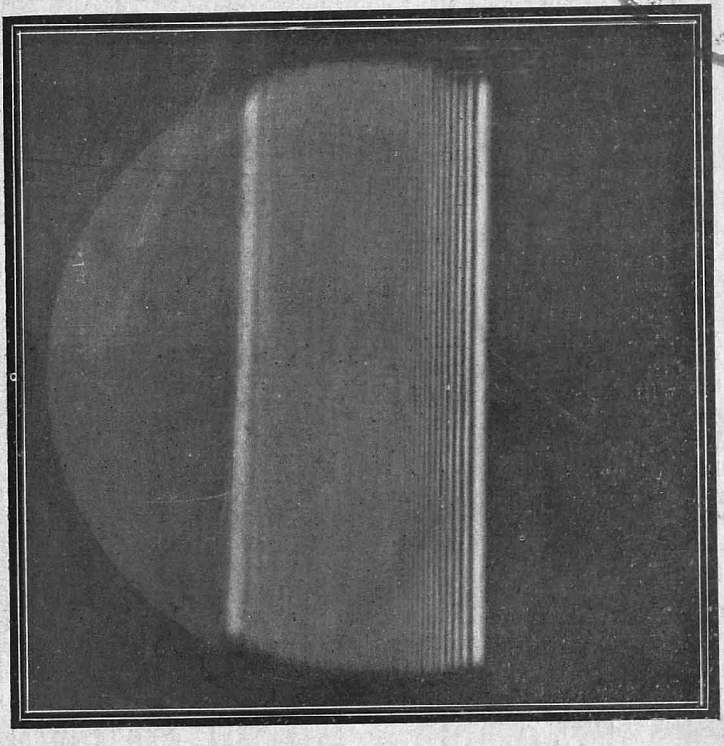


FIG. 1

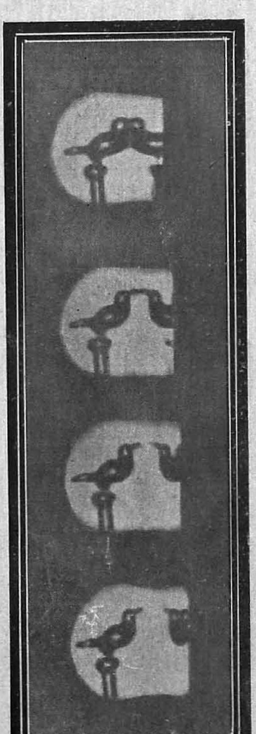


FIG. 2

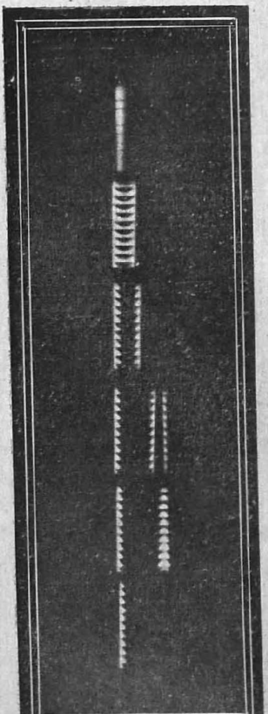


FIG. 3

of the phenomena actually noticed in mirages on a large scale in the open-air. The illustrations accompanying the present article, Figs. 1, 2 and 3 are also taken from the paper under reference.

Briefly stated, the analysis indicates that when $\mu_1 \cos \phi_1$ is greater than μ_2, the light incident on the front face of the slab is returned from its interior with full intensity and in a direction justifying its description as a regularly reflected disturbance. The analysis also shows that within the slab, the intensity of illumination attains a large value in the vicinity of the limiting plane at which the refractive index μ_L has the value $\mu_1 \cos \phi_1$. The intensity falls off rapidly to small values in the rear of that plane, while in front of it, the intensity diminishes gradually at the same time exhibiting a series of maxima and minima of which the separation falls off progressively, finally approaching a constant value.

The situation described above closely resembles the effects well known to all students of optics which are observed in the vicinity of *caustics*. Accordingly, we may in the present case state that the incident and reflected disturbances join up and form a cusped wave-front; the cusp rests upon the limiting plane referred to above and the entire wave-front moves along that plane in a direction parallel to the plane of incidence. At and near the cusp, the two branches of the wave-front are sensibly parallel to each other, besides being normal to the plane of the stratifications. Further out, however, they separate and diverge, ultimately becoming normal to the incident and reflected rays in the sense of geometric optics. The progressive diminution in the spacing of the successive maxima of illumination is readily understood on this basis.

6. Observation of the Caustic and Accompanying Interferences

Mirages can be produced and observed on a small scale in the laboratory as has been shown by R. W. Wood and others. In the arrangements generally used, the hot plate above which the mirages are observed is held horizontally. Though in some respects this arrangement is very convenient, it is not useful for critical studies owing to the thermal instability of the air above a heated plate. This difficulty is minimised by holding the heated plate edgewise so that its surface is vertical while the length remains horizontal. The object studied is an illuminated slit kept at some distance from the heated plate parallel to its surface. The light diverging from the slit is rendered parallel by a collimating lens and then allowed to fall obliquely on the heated plate. The beam is allowed to cover the whole length of the plate, the angle of incidence being adjustable by a lateral movement of the illuminated slit.

Photographs of the field of illumination in the rear of the hot plate could be secured with these arrangements and with very short exposures, provided that sunlight is used to illuminate the slit. With such short exposures the effects of thermal instability are eliminated completely. A typical example of a photograph thus obtained is reproduced in Fig. 1 accompanying. The use of a red filter to monochromatise the light greatly improves the results obtained. It will be seen that the features indicated by the theory and other features as well are very clearly exhibited.

The field of illumination appearing in Fig. 1 consists of three parts. To the right of the bright caustic (*i.e.*, towards the heated surface) the field is dark, while to its left lies an illumined strip containing a large number of interference fringes whose separation narrows down to a constant value as we move away from the caustic; since the heated plate is necessarily of finite extension, the reflected part of the cusped wave-front does not extend to infinity but is terminated. This manifests itself in the field of view by the occurrence of a second edge to the left of which the intensity is considerably less (though not zero), the edge being bordered by some broad fringes.

7. Relation of the Mirage to the Cusped Wave-Front

The actual mirage is observed when the eye is kept at any point which lies on the bright strip of light lying to the left of the caustic, the eye being focussed on the plane containing the object, *i.e.*, at infinity. It is to be expected that two images would then be seen whose positions lie respectively along the directions of the normals drawn from the eye to the two branches of the cusped wave-front leaving the nearer edge of the plate. In fact, the fringes

observed to the left of the caustic in Fig. 1 may be regarded as due to the interference between the light from two such virtual sources, the progressive narrowing of the spacing of the fringes to the left of the caustic corresponding to the increasing separation of the sources. That the separation of the two images observed depends on the position of the eye or aperture through which the phenomenon is viewed is illustrated in Fig. 3. In order to make the nature of the image evident, the serrated edge of a hacksaw blade has been used to form one of the edges of the slit. An aperture was kept before the camera lens and the succession of photographs exhibit the alteration in the phenomena as the aperture is gradually moved to the left. When the aperture is on the bright caustic, a single image is seen, while as it is moved to the left this separates into two images, one of which is a direct (or more properly, a refracted) image, the second being an inverted reflected image, the separation between the two gradually increasing. A remarkable feature of the sequence of phenomena illustrated in Fig. 3 is the occurrence of a third erect image close to the reflected image in the fourth and fifth photographs of the sequence; this image starts developing when the aperture has been moved towards the outer edge of the central illumined strip (where broad fringes start appearing in Fig. 1) and becomes coincident with the reflected image when the aperture is exactly at the edge mentioned. As the aperture is moved further left, only the 'direct' image continues to be visible, as is shown in the last photograph of the sequence; this is to be expected since the reflected part of the cusped wave-front is no longer received through the aperture.

The third image mentioned above could be cut off by inserting an opaque screen near the farther end of the heated plate and adjusting it so that its edge protrudes a little beyond the surface of the plate. This shows that the image is due to the ordinary refraction of rays directly entering the edge of the heated layer at the farther end of the plate. The main features of the path of such rays may be deduced from the experimental observations described in the previous paragraph. The terminus of the reflected part of the cusped wave-front corresponds to certain limiting rays entering the region at the farther end of the plate. Rays which are able to enter the edge of the heated layer at a closer distance to the plate than these limiting rays proceed a longer distance before emerging from the heated stratum and also suffer a larger deviation. These rays give rise to the erect third image; in fact, the second refracted wave-front, obtained by drawing the surfaces orthogonal to these rays, meets the termination of the refracted part of the wave-front so as to form a second cusp.

Till now we have dealt mainly with the case when the distant object is of negligible angular dimensions. When an object of finite angular dimensions is used, the point on the image which corresponds to any particular point on the object is to be determined as before for each setting of the eye. In this case there will be a distortion of the images because the position of the limiting layer as well as the inclinations of the cusped wave-fronts corresponding to any particular point on the object varies with the position of the object-point. Figure 2 shows the photographs taken using as the object a small model of a bird made of glass. This was placed near the focal point of the collimating lens. The sequence of photographs show the variation in the appearance of the phenomena as the eye is moved away from the plane of the plate. The appearance of a third erect image in addition to the usual reflected image may be discerned in the last two photographs of the sequence. It is worthy of note that a third erect image adjoining the inverted image which is the principal feature of the mirage has actually been noticed by various observers in mirages as seen in a large scale in the open air. It was noticed, for example, by Hiller in his study of mirages of the inferior type produced by a long vertical wall which had been warmed by the sun's rays. Hiller's photograph showing this effect is reproduced in Pernter's treatise already mentioned and also elsewhere. The third erect image is a conspicuous feature in Vince's drawings of ships at sea exhibiting the phenomenon of the superior mirage. It may safely be presumed that the explanation of its appearance is analogous to that of the effect noticed in the laboratory experiments and illustrated in Figs. 2 and 3 above.

Reprinted from "The Proceedings of the Indian Academy of Sciences",
Vol. L, No. 4, Sec. A, 1959

THE RAMAN AND INFRA-RED SPECTRA OF SOME SOLID HYDROXIDES

Part I. Raman Effect Data

BY D. KRISHNAMURTI, F.A.SC.
(*Memoir No.* 119 *from the Raman Research Institute, Bangalore*–6)

Received August 27, 1959

1. INTRODUCTION

A STUDY of the Raman spectra of the hydroxides of the alkali and alkaline earth group of metals is inherently of interest in view of the simplicity of their chemical composition. There have been comparatively few investigations regarding them and the results reported refer only to the frequency shifts in the $3\,\mu$ region. This paucity of data is attributable to the experimental difficulties of the subject; these substances are available only in the form of small crystals, polycrystalline masses or fine powders, and when the radiations of the mercury arc in the visible region are used for the excitation of the Raman spectra, the parasitic illumination entering the spectrograph is so strong as to obliterate any low-frequency shifts that might be observable in the vicinity of the exciting radiation. When the technique of exciting the Raman effect by the $\lambda\,2536\cdot5$ resonance radiation of mercury is employed, these difficulties are minimized and it becomes possible to record the low-frequency spectra at least in those cases where the material exhibits some measure of crystallinity and transparency. Since several of the hydroxides are also available as hydrates, it is of interest to study the Raman spectra of these hydrates as well. The first part of the present memoir deals with the facts of observation emerging from the experimental studies made with these substances.

Since the crystal structures of most of these substances have been determined already by X-ray diffraction methods and sometimes even through neutron-diffraction studies, the details of the spectra recorded are amenable to interpretation by the application of group theoretical methods and the second part of the memoir concerns itself with this. As all these substances crystallize in the holohedral class of their respective crystal systems, the rule of mutual exclusion is operative and hence the results of the Raman effect studies are complementary to those derived from infra-red studies and the latter do not therefore render the former superfluous.

The few infra-red investigations so far reported on these substances have all been restricted to the high-frequency region of 3 μ to 15 μ, evidently because of the experimental difficulties in the study of the far-infra-red region of the spectrum. That the present Raman effect data also afford a satisfactory explanation of the complex infra-red absorption spectra, reported in some cases in the 3 μ region, will emerge in the third part of the memoir wherein we shall consider the infra-red data in relation to the Raman spectra of these substances.

2. EXPERIMENTAL TECHNIQUES

The spectra were recorded with a Hilger medium quartz spectrograph and in some cases with a Hilger Littrow spectrograph of higher resolving power. The general details regarding the experimental set-up and about the water-cooled magnet-controlled quartz mercury arc emitting the intense λ 2536·5 resonance radiation which was used for the excitation of the Raman effect have already been described in an earlier paper by the author (1957). It is pertinent to emphasize here that the results to be reported have been achieved chiefly because of the advantages inherent in the use of λ 2536·5 excitation, *viz.*, the intense scattering power in the ultra-violet region and the effective suppression by a filter of mercury vapour inside the spectrograph of the parasitic resonance radiation reaching the slit of the spectrograph, thereby enabling us to record the low-frequency spectrum clearly even in the cases of coarse crystals and polycrystalline masses. The large reactivity of these substances with the atmospheric moisture and carbon dioxide necessitated that they be sealed up in quartz tubes before recording their spectra.

With a slit width of 0·05 mm. exposures of the order of ten hours were found to be usually adequate to record the spectra with reasonable intensity and contrast. With increased exposures only the weaker mercury lines and the continuous background accompanying the spectrum of the mercury arc became prominent and any weak Raman lines recorded were masked by the general background intensity. Sometimes a few weak mercury lines appeared on the anti-Stokes side of the spectrum and care has been taken not to identify any of them as frequency shifts unless Raman lines of the same frequency shifts were present on the Stokes side as well. As a particular case of this, mention may be made of a group of lines noticed on the anti-Stokes side of the spectrum of $Ba(OH)_2 \cdot 8H_2O$, and which were not identified as Raman frequency shifts.

The following is a brief description of the several substances whose spectra were recorded in the present investigation:

(i) *Lithium hydroxide* (*anhydrous*).—The sample, made available through the kindness of a friend, was in the form of porous granules and exhibited no visible crystallinity. The sample actually contained a good proportion of $LiOH.H_2O$ as was revealed by the spectrum which exhibited the Raman lines due to pure $LiOH.H_2O$ also.

(ii) *Lithium hydroxide monohydrate*.—The material in the form of small lumps containing minute cyrstals was kindly provided by the Head of the Department of General Chemistry, Indian Institute of Science, Bangalore.

(iii) *Sodium hydroxide* (*anhydrous*).—The sodium hydroxide used in the study was the translucent polycrystalline mass commercially available in the form of sticks and manufactured by British Drug Houses, Ltd., and was of analytical reagent quality.

(iv) *Sodium hydroxide monohydrate*.—The monohydrated compound of NaOH was prepared by cooling a hot concentrated solution of the substance. The monohydrate is formed between 12·3° C. and 64·3° C. (Mellor, 1946).

(v) *Potassium hydroxide* (*anhydrous*).—In the case of KOH, the commercially available substance was found to be of the hydrated variety. Therefore, the pure substance (manufactured by the firm Merck & Co.) was heated in a nickel crucible in an attempt to drive out the water. The substance thus prepared was a white mass of fine texture. In spite of the heating, however, it was found that the spectrum continued to exhibit the hydroxyl frequency shift due to the hydrate by the side of that due to the anhydrous substance.

(vi) *Potassium hydroxide monohydrate*.—Near room temperature there are two hydrates of KOH, *viz*., the dihydrate and the monohydrate. The dihydrate is stated to melt at about 35° C. or slightly below and pass over into the monohydrate which is stable up to 143° C. (Mellor, 1946). The substance used in the present investigation was prepared by cooling a highly concentrated hot solution of KOH and the stability of this substance was checked up to 50° C. and was identified as the monohydrate.

(vii) *Calcium hydroxide*.—This is normally available only in the form of a fine powder. Crystals of this substance were grown by slowly evaporating a large volume of a saturated solution of $Ca(OH)_2$ at about 80° C. inside a thermostat. Owing to the very low solubility of the substance, the crystals obtained were quite small, being only a fraction of a millimetre in size. The material thus prepared contained a small proportion of calcium carbonate formed by the action of atmospheric carbon dioxide, as was revealed by the presence (faintly) of the principal frequency shift of the CO_3 group in the spectrum recorded.

(viii) *Magnesium hydroxide.*—The compound in the form of a colloidally fine powder was prepared by precipitation and was kindly made available by the Head of the Department of General Chemistry, Indian Institute of Science, Bangalore.

(ix) *Strontium hydroxide octahydrate.*—The substance in the form of small crystals, manufactured by Merck & Co., was used directly as such to record the spectra. The tetragonal crystal forms and the uniaxial nature of the crystal served to confirm their identification as the common octahydrate of $Sr(OH)_2$.

(x) *Barium hydroxide octahydrate.*—Small crystals of this substance (manufactured by Johnson & Sons) were used directly as such to record the spectra. That they were the common octahydrate of $Ba(OH)_2$ was confirmed by their biaxial nature and by the small optic axial angle exhibited by them (Winchell, *Microscopic Characters of Artificial Minerals,* pp. 187).

(xi) *Aluminium hydroxide.*—'Dry' aluminium hydroxide manufactured by British Drug Houses Ltd., in the form of fine powder, was used to obtain the spectra.

3. RESULTS

The frequency shifts observed in the case of the different substances are shown in Table I. The following are the principal features exhibited by the spectra of the several hydroxides. The observed frequency shifts group themselves and naturally fall into two specific regions, *viz.*, (*a*) a region of low-frequencies from about 75 $cm.^{-1}$ to 400 $cm.^{-1}$ and (*b*) a region between 3300 to 3700 $cm.^{-1}$ and which can readily be recognized as respectively the range of lattice frequencies and of hydroxyl frequencies. It is noteworthy that no frequency shifts have been observed in the region intermediate to the above ranges. The extremely fine texture and near opacity of some of the materials used in the study, *viz.*, LiOH, KOH, $Mg(OH)_2$ and $Al(OH)_3$, and the intense parasitic illumination consequent thereon reaching the slit of the spectrograph made it impossible to observe any low-frequency shifts in their cases. The spectra recorded have been grouped to form two Plates (II and III) and reproduce respectively the low and the high frequency regions.

In the majority of cases the high-frequency region exhibits only a single intense and sharp frequency shift contrary to our experience of the broad and diffuse bands observed with water and water of crystallization in the case of hydrated crystals. It is a surprising fact that in the cases of $LiOH \cdot H_2O$, $NaOH \cdot H_2O$ and $KOH \cdot H_2O$ the Raman bands due to water of crystallization

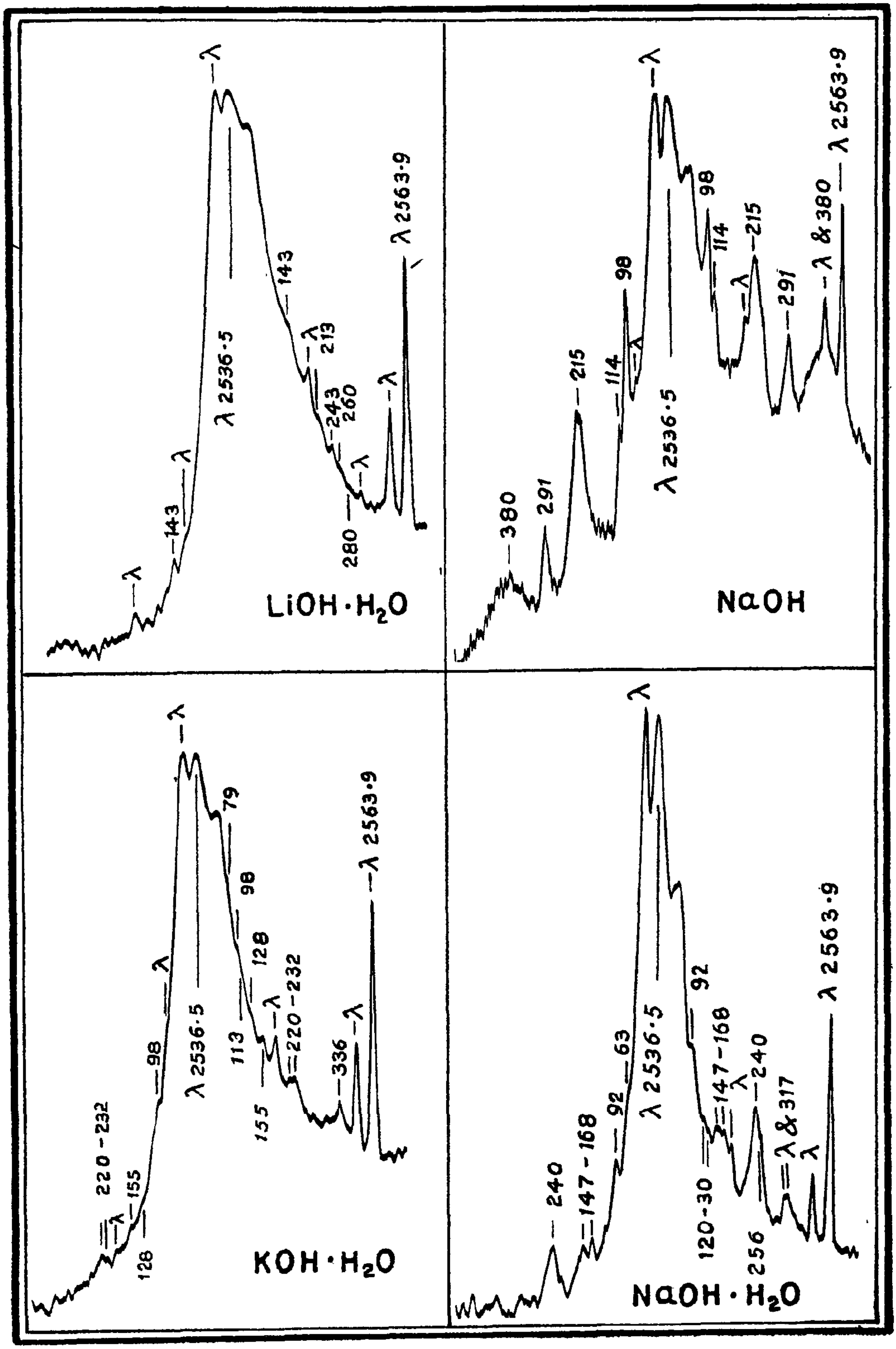


FIG. 1

TABLE I

Observed Raman frequency shifts in cm.$^{-1}$

Substance	Lattice frequencies	Internal frequencies
LiOH ..	..	3664 (*v.s.*)
$LiOH \cdot H_2O$ ..	143 (*s.*), 213 (*m.*), 243 (*m.*), 260 (*w.*), 280 (*w.*)	3563 (*v.s.*)
NaOH ..	98 (*v.s.*), 114 (*m.*), 215 (*v.s.*), 291 (*s.*), 380 (*w.b.*)	3632 (*v.s.*)
$NaOH \cdot H_2O$ ..	63 (*s.*), 92 (*s.*), 120 (*w.*), 130 (*w.*), 147 (*m.*), 168 (*m.*), 240 (*v.s.*), 256 (*w.*), 317 (*m.*)	3568 (*v.s.*)
KOH ..	..	3597 (*s.*)
$KOH \cdot H_2O$ ..	79 (*m.*), 98 (*s.*), 113 (*s.*), 128 (*s.*), 155 (*m.*), 220 (*m.*), 232 (*m.*), 336 (*m.*)	3501 (*v.s.*)
$Mg(OH)_2$ ..	..	3651 (*s.*)
$Ca(OH)_2$ ..	247 (*m.*), 282 (*m.*), 359 (*s.*)	3616 (*v.s.*)
$Sr(OH)_2 \cdot 8H_2O$	106 (*m.*), 132 (*s.*), 155 (*w.*), 173 (*m.*), 199 (*w.*), 230 (*v.s.*), 297 (*m.*), 328 (*w.*)	3302 (*m.b.*) 3393 (*m.b.*) 3488 (*v.s.*) 3536 (*w.*) 3593 (*w.*)
$Ba(OH)_2 \cdot 8H_2O$	106 (*s.*), 139 (*m.*), 156 (*m.*), 173 (*w.*), 205 (*v.s.*), 220 (*m.*), 276 (*m.*), 287 (*m.*), 328 (*w.*)	3335 (*w.b.*) 3481 (*v.s.*) 3529 (*s.*)
$Al(OH)_3$ ..	..	3369 (*s.*) 3431 (*m.*) 3521 (*s.*) 3583–3602 (*m.*)

s. = strong; *w.* = weak; *m.* = moderate intensity; *v.s.* = very strong; *b* = broad.

have not been recorded even though the hydroxyl frequency shift appears with large intensity. Even in the case of $Sr(OH)_2 \cdot 8H_2O$ and $Ba(OH)_2 \cdot 8H_2O$ the water bands appear only weakly in comparison with the large intensity of the shifts due to the hydroxide.

It may be noticed that the hydroxyl frequency shift (in anhydrous compounds) is generally higher in the case of substances with metallic ions of

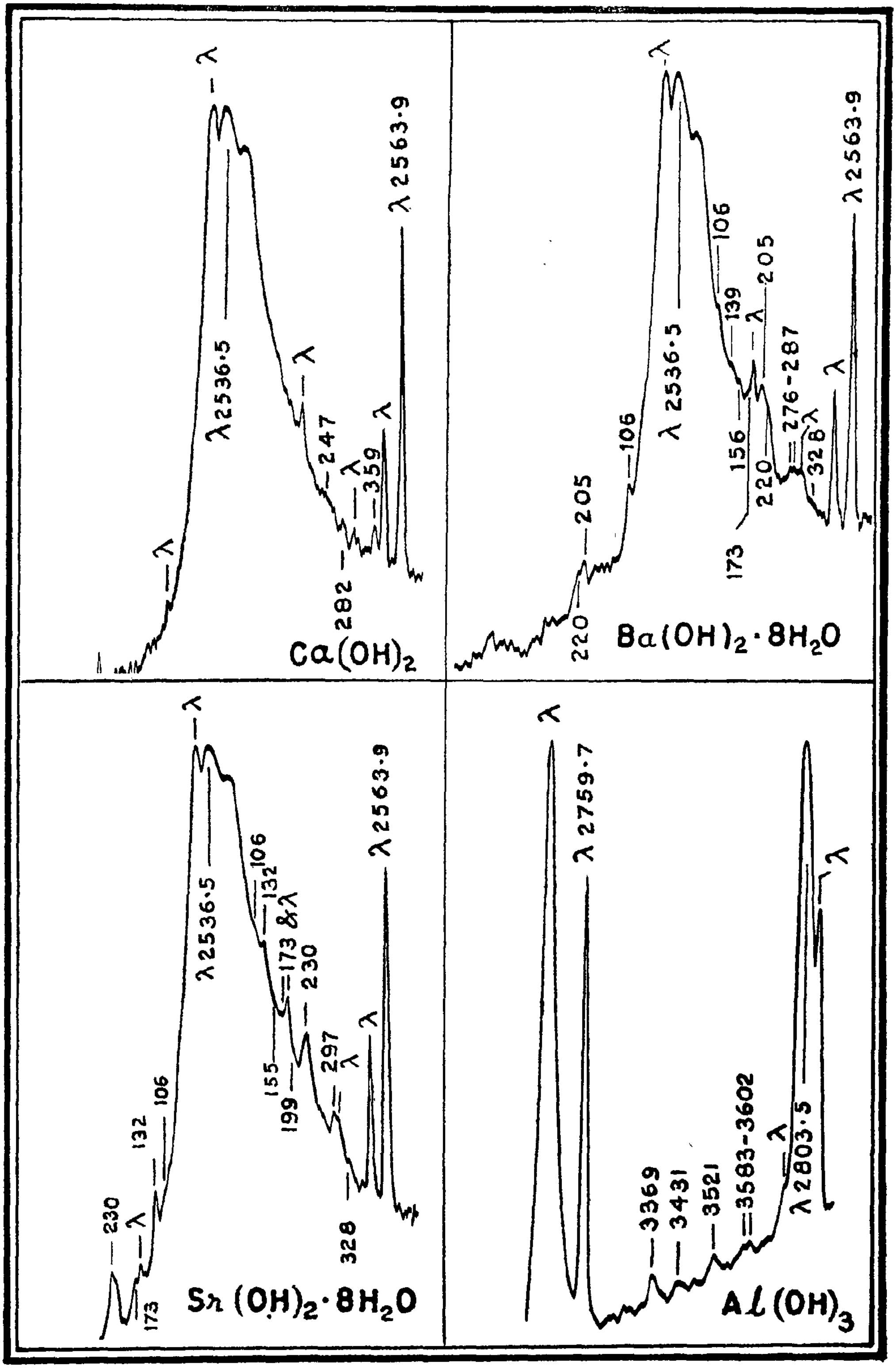

Ca(OH)2
λ2536·5
λ2563·9
247
359
282
Ba(OH)2·8H2O
λ2536·5
λ2563·9
106
139
205
156
220
276-287
328
173
Sr(OH)2·8H2O
λ2536·5
λ2563·9
106
132
173 & λ
230
155
199
297
328
Al(OH)3
λ2759·7
λ2803·5
3369
3431
3521
3583-3602

Fig. 2

lower atomic weight. The hydroxyl frequency shifts in the cases of the hydrated hydroxides also form a similar sequence and are considerably lower (and less sharp) than the frequency shifts observed with the corresponding anhydrous compounds.

The lattice frequency shifts in the several cases exhibit diverse features with regard to their number, intensity and sharpness, all of which evidently have their origin in the differences in their crystal structures. To a discussion of these facts we shall return in the next part of the memoir.

The frequency shift of 282 cm.$^{-1}$ observed with Ca $(OH)_2$ coincides with the value of one of the lattice frequencies of calcite. The spectrum of the slightly impure Ca $(OH)_2$, as already mentioned, showed weakly the principal frequency 1086 cm.$^{-1}$ of calcite. However, the 282 cm.$^{-1}$ shift in calcite is relatively weaker than the 1086 cm.$^{-1}$ shift and it is highly improbable that the impurity of $CaCO_3$ was responsible for the 282 cm.$^{-1}$ shift. The identification of the shift as genuinely due to the Ca $(OH)_2$ is vindicated by the infra-red data as we shall see in Part III.

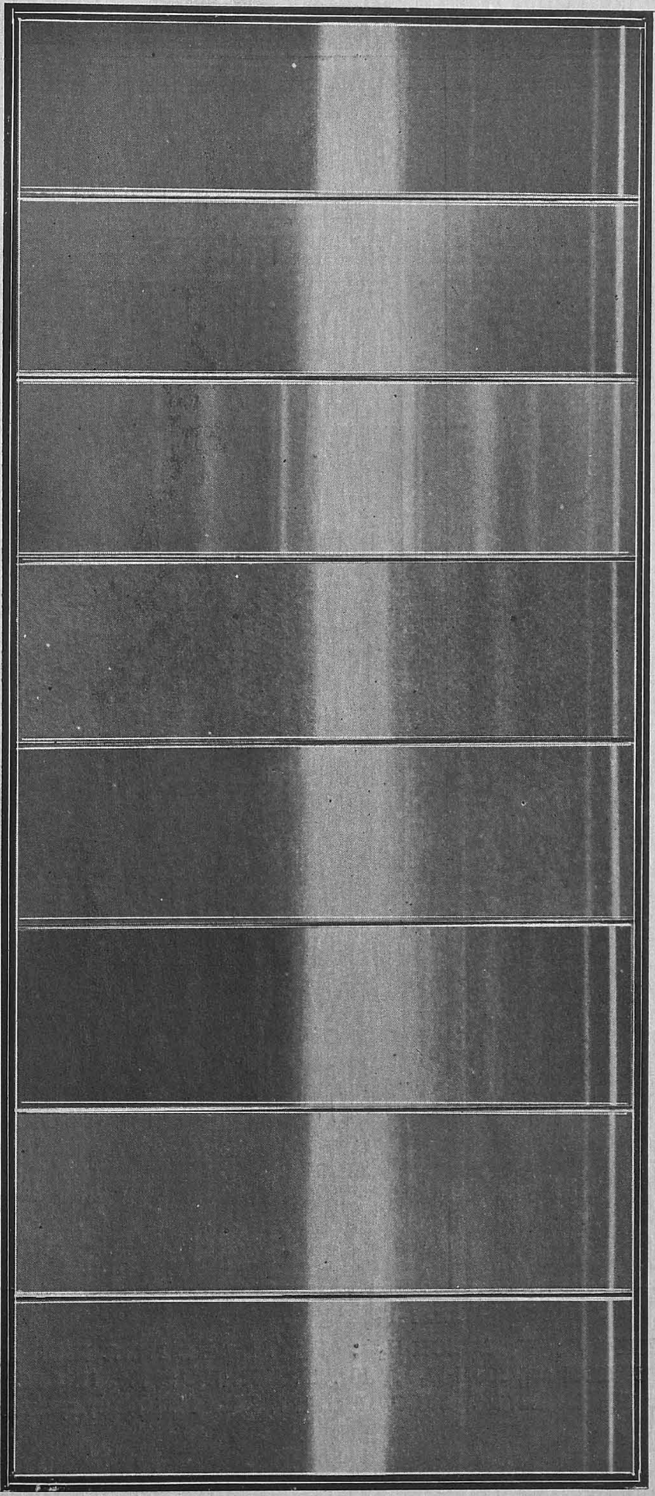
In order to exhibit several of the frequency shifts which were recorded rather faintly, microphotometer records of different cases (for the lattice frequency region) appear in the text. Figure 2 in the text includes also the microphotometer record of the spectrum of Al $(OH)_3$ in the hydroxyl frequency region since the spectrum recorded was too weak for reproduction.

Some of the frequency shifts shown in Table I, *viz.*, 3632 cm.$^{-1}$ in the case of NaOH, and 3302, 3393 and 3488 cm.$^{-1}$ in the case of Sr $(OH)_2 \cdot 8H_2O$ and 3481 and 3529 cm.$^{-1}$ in the case of Ba $(OH)_2 \cdot 8H_2O$ have been reported by the earlier investigators, with minor differences in the observed values (Krishnamurti, 1930; Theimer, 1950; Busing, 1955).

The author wishes to express his sincere thanks to Professor Sir C. V. Raman, F.R.S., N.L., for his keen interest in this investigation.

4. SUMMARY

The paper reports the results of an investigation on the Raman spectra of eleven solid hydroxides using the λ 2536·5 resonance radiation of mercury arc for the excitation of the Raman effect. The cases studied (most of them for the first time) consist of the hydroxides of Li, Na and K and their monohydrates, the hydroxides of Ca, Mg and Al, and the octahydrated hydroxides of Sr and Ba. The facts of observation are discussed in relation to the crystal structure and the infra-red absorption data of these crystals in two separate parts following this paper.



Hg arc

$LiOH \cdot H_2O$

NaOH

$NaOH \cdot H_2O$

$KOH \cdot H_2O$

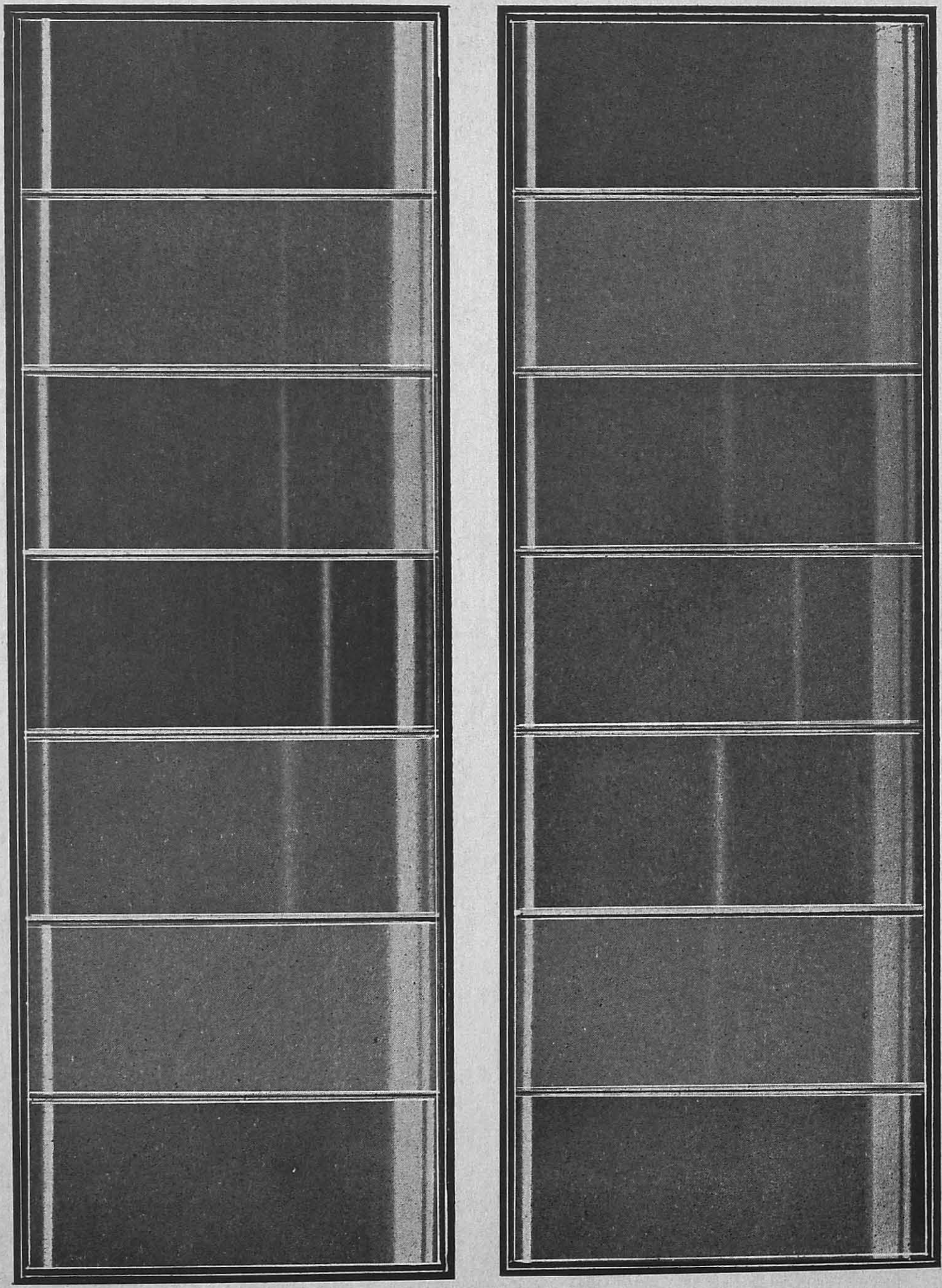
$Sr(OH)_2 \cdot 8H_2O$

$Ba(OH)_2 \cdot 8H_2O$

Hg arc

CONNEMARA PUB
4 JAN 196

(1) (2) (3) (4) (5) (6) (7)



(8) (9) (10) (11) (12) (13) (14)

FIGS. 1–14

Fig. 2. LiOH and LiOH·H_2O; Fig. 3. LiOH·H_2O; Fig. 4. NaOH; Fig. 5. NaOH·H_2O; Fig. 6. $Mg(OH)_2$; Fig. 9. KOH and KOH·H_2O; Fig. 10. KOH·H_2O; Fig. 11. $Ca(OH)_2$; Fig. 12. $Sr(OH)_2 \cdot 8H_2O$; Fig. 13. $Ba(OH)_2 \cdot 8H_2O$. Figs. 1, 7, 8 and 14. Mercury spectrum.

REFERENCES

Busing, W. R. .. *J. Chem. Phys.*, 1955, **23**, 933.

Krishnamurti, D. .. *Proc. Ind. Acad. Sci.*, 1957, **46 A**, 183.

Krishnamurti, P. .. *Ind. Jour. Phys.*, 1930, **5**, 651.

Mellor, J. W. .. *A Comprehensive Treatise on Inorganic and Theoretical Chemistry*, Longmans, Green & Co., 1946, **2**, pp. 502 *et seq.*

Theimer, O. .. *Mh. Chem.*, 1950, **81**, 301.

THE RAMAN AND INFRA-RED SPECTRA OF SOME SOLID HYDROXIDES

Part II. Correlation with Crystal Structure

BY D. KRISHNAMURTI, F.A.SC.

(*Memoir No.* 119 *from the Raman Research Institute, Bangalore*-6)

Received August 27, 1959

1. INTRODUCTION

THE application of conventional methods of group theory for elucidating the features observed in the Raman and infra-red spectra of crystals is well known and is dealt with in detail by Bhagavantam and Venkatarayudu in their treatise on the Theory of Groups. The group theoretical methods give the total number of vibrations appearing under each of the different symmetry classes and in addition it is usually possible to resolve this number into translatory, rotatory and internal vibrations. It had been pointed out by Bhagavantam[13] in his discussion of the case of crystalline Hg_2Cl_2 that in the case of crystals containing linear groups (like the hydroxyl ions figuring in our discussions) which possess only two degrees of rotational freedom, the usual formulæ used for resolving the total number of vibrations into the lattice and internal types of vibrations are inapplicable. However, since the symmetry properties and selection rules of the different irreducible representations of the various point groups have all been worked out and are readily available for reference (*vide e.g.*, *Molecular Vibrations* by Wilson *et al.*) it is possible from simple considerations of symmetry to picture the nature of the individual symmetry modes in simple cases. In the following we shall adopt this procedure of identification wherever feasible.

Crystal structure data naturally play an important role in the explanation of the observed facts. The relevant data are summarized in Table I. A list of references appears at the end of the paper and the numbers appearing in the first column of Table I refer to the concerned papers. In column 4 of the same table are shown the distances between the oxygens of the hydroxyl groups and in the case of the hydrates the figures marked with an asterisk denote the distances between the oxygens of the hydroxyls and the oxygens of the water molecules.

TABLE I

Details of crystal structure

Substance	Crystal system	Space group and number of units in primitive cell	O–O distance in A.U.	Characteristic features of the structure
LiOH[1] ..	Tetra-gonal	D^7_{4h} (2)	3·6	Layer structure normal to *c*. Each Li is surrounded by four hydroxyls nearly tetrahedrally. Each hydroxyl has four Li neighbours on one side and four hydroxyls of the next layer on the other side. Hydroxyls are parallel to *c*-axis. Site symmetry of Li: D_{2d}; of OH: C_{4v}.
$LiOH \cdot H_2O$[2]	Mono-clinic	C^3_{2h} (2)	2·99 3·14 to 3·34*	Chain structure. Each Li surrounded nearly tetrahedrally by two hydroxyls and two water molecules. Site symmetry of oxygens of OH: σ_h; of Li: C_2; of O (H_2O): C_2.
NaOH[3–4] KOH[3]	Ortho-rhombic	D^{17}_{2h} (2) do.	3·40 3·26	Layer structure normal to *c*. Each Na is surrounded by five OH groups. OH ions are parallel to *c*. Na and O approximately in the *ab* plane. Characteristic double layers of Na and OH normal to *c*. Site symmetry for both Na and OH: C_{2v}. Similarly for KOH.
$NaOH \cdot H_2O$[5]	Ortho-rhombic	D^{15}_{2h} (8)	3·5 to 4·5 2·64 to 3·11*	Layer structure normal to *b*. No special positions of symmetry. Each Na surrounded by three hydroxyls and one H_2O nearly tetrahedrally. Each OH surrounded by 3 Na's and 2 H_2O's, and 3 OH's very much farther out than Na and H_2O.

TABLE I (*Contd.*)

Substance	Crystal system	Space group and number of units in primitive cell	O–O distance in A.U.	Characteristic features of the structure
$KOH \cdot H_2O$[6] ..	Mono-clinic	C^5_{2h} (4)	2·63*	Layer structure parallel to (100). No special positions of symmetry. Oxygens ($OH—OH_2—OH$) form zig-zag chains parallel to *c*.
$Mg(OH)_2$[7] .. $Ca(OH)_2$[8–9] ..	Trigonal do.	D^3_{3d} (1) do.	3·22 3·33	Layer structure normal to *c*. Hydroxyls parallel to *c*. Site symmetry of Mg: D_{3d}; of OH: C_{3v}. Each Mg is surrounded by six hydroxyls. Each OH is surrounded by three Mg's on one side and on the other side by three OH's of the next layer. Similarly for $Ca(OH)_2$.
$Sr(OH)_2 \cdot 8H_2O$[10]	Tetra-gonal	D^8_{4h} (4)	2·9 2·7*	Site symmetry of Sr: D_2; of oxygens of OH: C_4. The rest are in general positions. Each OH surrounded octahedrally by 4 H_2O's and 2 OH's.
$Ba(OH)_2 \cdot 8H_2O$[11]	Mono-clinic	C^5_{2h} (4)		Atomic positions not determined. But all atoms are in general positions.
$Al(OH)_3$[12] ..	Mono-clinic	C^5_{2h} (8)	2·75 to 3·46	All atoms in general positions. Distorted version of $Mg(OH)_2$ structure, the double layers of hydroxyls being parallel to (001). But only two-thirds of octahedral holes between hydroxyl layers are filled by Al. Owing to the distorted nature not all hydroxyls are parallel to *c* as in $Mg(OH)_2$.

* See text above for explanatory remarks.

2. The High Frequency Region of the Spectrum

The following facts regarding the crystal structures of the hydroxides enable us to readily account for the general features observed in the high frequency region of the spectrum. All the crystal structures shown in Table I belong to the centro-symmetric classes and the hydroxyl groups in the unit cell are separable into two halves related to each other by the centre of inversion. In many of the cases [*viz.*, LiOH, $LiOH \cdot H_2O$, NaOH, KOH, $Mg(OH)_2$ and $Ca(OH)_2$] the unit cell contains only two hydroxyl groups which are on either side of the centre of inversion. As a consequence there arise two coupled hydroxyl stretching vibrations which are respectively symmetric and antisymmetric with respect to the centre of inversion. In particular, in the cases of LiOH, NaOH, KOH, $Mg(OH)_2$ and $Ca(OH)_2$ where the hydroxyl groups occupy special positions of symmetry their symmetric vibrations actually belong to the totally symmetric class. In the case of $LiOH \cdot H_2O$ the positions of the hydrogens of the hydroxyls are not known. However, since it is remarked by Levy and Peterson[14] that the hydroxyl ion in $LiOH \cdot H_2O$ does not take part in hydrogen bonding, it appears highly probable that the hydrogens of the hydroxyls are situated on the symmetry plane, and the symmetric vibrations of the hydroxyls in this case also can easily be seen to belong to the class A_g of C_{2h}. The symmetric vibrations in all these cases are active only in the Raman effect. The vibrations appearing in the infra-red are antisymmetric and Raman inactive. The observation of a single frequency shift in each one of the above cases is hence in accordance with theory, and these shifts are identified with the internal vibrations of the hydroxyls belonging to the totally symmetric class.

With regard to substances which contain more units in the primitive cell, the coupled internal vibrations of the hydroxyls can be divided into two groups of equal number of vibrations which are characterized by the fact of their being symmetric or antisymmetric with respect to the centre of inversion. With the exception of $Sr(OH)_2 \cdot 8H_2O$, in all the remaining cases wherein the atoms are situated in general positions, *i.e.*, in the case of $NaOH \cdot H_2O$, $KOH \cdot H_2O$, $Ba(OH)_2 \cdot 8H_2O$ and $Al(OH)_3$, all vibrations which are symmetric with respect to the centre of inversion are Raman active. Hence, the spectra of $NaOH \cdot H_2O$, $KOH \cdot H_2O$, $Ba(OH)_2 \cdot 8H_2O$ and $Al(OH)_3$ should according to theory exhibit respectively 4, 2, 4 and 12 components due to the internal vibrations of the hydroxyls. However, it is obvious that only the modes in which the internal vibrations of all the hydroxyls take place in identical phase (*i.e.*, totally symmetric class) would be intensely Raman active, the others being relatively weak. Also, if the hydroxyl groups are far removed from each other as in $NaOH \cdot H_2O$ the coupling between the

hydroxyls will be too weak for the differences in the phases of adjacent groups to influence the frequency shift observably. In particular, in $NaOH \cdot H_2O$ the eight units of the primitive cell are distributed equally in two adjacent layers which are quite far apart, the spacing being 5·9 A.U. In these circumstances, the appearance, in the spectra, of fewer components than that predicted by theory may be ascribed to either the low intensity of some of the components and (or) to their falling too close to the totally symmetric frequency to be resolved from it.

In view of the foregoing remarks the frequency shifts observed at 3568 $cm.^{-1}$ in $NaOH \cdot H_2O$ and at 3501 in $KOH \cdot H_2O$ may be assigned to the totally symmetric class. The appearance of two strong components well separated from each other at 3481 and 3529 $cm.^{-1}$ in the case of $Ba(OH)_2 \cdot 8H_2O$ is indicative of considerable coupling between the hydroxyls of the unit cell; the former which is more intense evidently belongs to the totally symmetric class. In the case of $Al(OH)_3$ five components, some of them rather broad, have been observed at 3369, 3431, 3521, 3583 and 3602 $cm.^{-1}$ It is probable that some of them are unresolved groups of lines. The distribution of the observed frequencies over a considerable range is due to the large differences in the distances between the several hydroxyls in the unit cell and correspondingly varying degrees of interactions between them.

In the case of $Sr(OH)_2 \cdot 8H_2O$ since there are eight hydroxyls per unit cell, there are, in all, four coupled internal vibrations symmetric with respect to the centre of inversion. Unlike the cases discussed above, the group D_{4h} to which $Sr(OH)_2 \cdot 8H_2O$ belongs contains an irreducible representation A_{2g} symmetric with respect to the centre of inversion and which is Raman inactive. Since the positions of the hydrogen atoms are not known, it is not possible to determine the exact number of components active in the Raman effect. Three components have been observed at 3488, 3536 and 3593 $cm.^{-1}$; the intense line at 3488 $cm.^{-1}$ most probably belongs to the totally symmetric class.

In the case of hydrated hydroxides, the spectrum should in addition exhibit frequency shifts due to the water molecules. According to theory the Raman spectrum of water vapour should exhibit three frequency shifts. Only a single strong frequency shift (of the symmetric class) at 3655 $cm.^{-1}$ had been so far observed, the two other vibrations being presumably too weak (*vide Infra-red and Raman Spectra* by Herzberg). However, they have been observed in the infra-red spectrum at 1595 and 3756 $cm.^{-1}$ Considerations similar to that put forward for the explanation of the hydroxyl

frequencies indicate that corresponding to the 3655 cm.$^{-1}$ shift of water vapour, the spectra of $LiOH \cdot H_2O$, $NaOH \cdot H_2O$, $KOH \cdot H_2O$ and $Ba(OH)_2 \cdot 8H_2O$ should exhibit respectively 1, 4, 2 and 16 components which arise due to the coupling of the vibrations and which take place symmetrically with respect to the centre of inversion. In the case of $Sr(OH)_2 \cdot 8H_2O$, 16 components are possible. But, the actual number that should be expected to appear in Raman effect cannot be determined in the absence of any data regarding the positions of the hydrogens. The fact that no shifts due to water are observed with $LiOH \cdot H_2O$, $NaOH \cdot H_2O$ and $KOH \cdot H_2O$ is rather surprising and it requires further experimental study with clear and large crystals to observe them. The broad bands at 3302 and 3393 in $Sr(OH)_2 \cdot 8H_2O$ and 3335 cm.$^{-1}$ in $Ba(OH)_2 \cdot 8H_2O$ are evidently due to the large number of shifts appearing unresolved from each other. The explanation of the diminution of these frequencies from the value of 3655 cm.$^{-1}$ for water vapour, we shall consider in the following section.

3. Hydroxyl Bonding Effects

The large variations in the hydroxyl frequency shifts observed in the different cases are explicable as arising due to the interactions between the different hydroxyls (and the water molecules as well in the hydrates) which are contiguous to each other. Bernal and Megaw[15] have discussed in detail the nature of the interactions between hydroxyl groups—called by them the hydroxyl bond to distinguish it from the hydrogen bond. In the case of interactions between hydroxyl groups the distances between the groups are comparatively large (ranging from 2·75 to 3·6 A.U.). (In the case of hydrogen bonding, the oxygen atoms approach to within 2·55 A.U. of each other and the hydrogen linking them is not preferentially attached to either of them.)

The studies on the structure of $Al(OH)_3$ and other hydroxides have led Bernal and Megaw (*loc. cit.*) to conclude that there exists an attractive force between the hydroxyls. The hydroxyl ion can be pictured to be made up of an O^{--} and a proton H^+ and as a result the negative charge of one oxygen can attract the positive charge of a neighbouring hydrogen to form the hydroxyl bond. It follows that the attraction will be greater at shorter O–O distances and will result in the displacement of the hydrogen atom from its normal position near the oxygen to which it belongs and as a consequence a diminution of the hydroxyl frequency would result. However, this is only a simplified picture of the actual state of affairs and, in general, the effects of all the surrounding neighbours will have to be taken into account. Especially in cases like LiOH, NaOH, KOH, $Ca(OH)_2$ and $Mg(OH)_2$ where the hydroxyl groups are far apart and are orientated parallel to each other and normal to

the layers, no large diminution of the frequency can be expected to occur. Even so, the fact that the hydroxyl frequency shifts in LiOH, NaOH and KOH (3664, 3632 and 3597 cm.$^{-1}$) follow the same sequence as the O–O distances in them (3·6, 3·4 and 3·26 A. U.) is significant. In the case of Ca $(OH)_2$ and Mg $(OH)_2$ the O—O distances are respectively 3·33 and 3·22 A.U. Their respective hydroxyl shifts 3616 and 3651 cm.$^{-1}$ fall in the region approximately near that of NaOH; however, it is surprising that the frequency shift in Mg $(OH)_2$ should be higher than that in the case of Ca $(OH)_2$.

In the case of the hydrates additional interactions arise between the hydroxyl ions and the water molecules contiguous to them. The progressive diminution of the frequency shifts in $LiOH \cdot H_2O$, $NaOH \cdot H_2O$, $KOH \cdot H_2O$ and $Sr(OH)_2 \cdot 8H_2O$ is evidently due to the diminishing O–O distances in these cases and the correspondingly large interactions between the hydroxyls and the water molecules contiguous to each other. The diminution of the frequency shifts of the water bands in $Sr(OH)_2 \cdot 8H_2O$ and $Ba(OH)_2 \cdot 8H_2O$ (from the normal value of 3655 cm.$^{-1}$ observed for water vapour) is also presumably due to analogous causes.

4. The Low-Frequency Region of the Spectrum

In the present section, we shall discuss the cases for which lattice frequency shifts are reported.

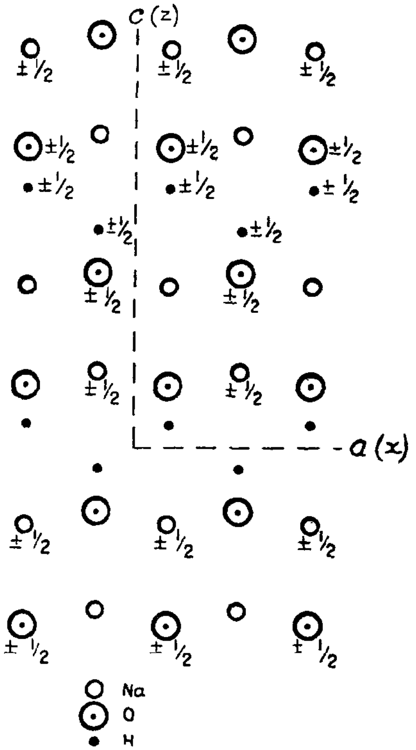
(i) *Lithium hydroxide monohydrate*.—The total number of vibrations to be expected in the Raman effect according to theory is 18. Amongst these are included the single symmetric internal vibration of the hydroxyls and three others due to the water present in the structure. Hence, fourteen vibrations should appear in the low-frequency region. However, it should be remarked that since $LiOH \cdot H_2O$ dehydrates easily to become LiOH when heated in an atmosphere of hydrogen below 140° C. (Partington, 1958),[16] the binding of the water molecules in the lattice is weak and the translational and rotational frequencies due to them will be rather low. It had not been possible to observe them. On the other hand, the binding between Li and OH should be quite strong as is evidenced by the high melting point of LiOH (450° C.). If we ignore the water molecules, there should appear in the Raman effect eight lattice frequency shifts, of which only five could be observed.

(ii) *Sodium hydroxide*.—The group characters and selection rules of the various irreducible representations of the group D_{2h}, to which NaOH belongs, are reproduced in Table II following. Figure 1 in the text illustrates the structure of NaOH.

TABLE II

D_{2h}	E	$C_2(z)$	$C_2(y)$	$C_2(x)$	i	$\sigma(xy)$	$\sigma(zx)$	$\sigma(yz)$	Activity
A_g ..	1	1	1	1	1	1	1	1	R.E. only
B_{1g} ..	1	1	−1	−1	1	1	−1	−1	R.E. only
B_{2g} ..	1	−1	1	−1	1	−1	1	−1	R.E. only
B_{3g} ..	1	−1	−1	1	1	−1	−1	1	R.E. only
A_u ..	1	1	1	1	−1	−1	−1	−1	Inactive in both
B_{1u} ..	1	1	−1	−1	−1	−1	1	1	I.R. only
B_{2u} ..	1	−1	1	−1	−1	1	−1	1	I.R. only
B_{3u} ..	1	−1	−1	1	−1	1	1	−1	I.R. only
U_R ..	6	6	0	0	0	0	6	6	

U_R = number of atoms invariant under the group operation.
x, y, z correspond to the crystallographic axes a, b and c respectively.



TEXT-FIG. 1. Projection of the structure of NaOH in a plane normal to the b-axis. ± ½ indicate that the atoms are at the levels ± ½b. The remaining atoms are in the plane passing through the origin.

TABLE III

Class	n_i	T	T′	R	n_i'	Activity
A_g ..	3	0	2	0	1	R.E.
B_{1g} ..	0	0	0	0	0	R.E.
B_{2g} ..	3	0	2	1	0	R.E.
B_{3g} ..	3	0	2	1	0	R.E.
A_u ..	0	0	0	0	0	Inactive
B_{1u} ..	3	1	1	0	1	I.R.
B_{2u} ..	3	1	1	1	0	I.R.
B_{3u} ..	3	1	1	1	0	I.R.

n_i = Total number of vibrations; T = pure translations; T′ = Translational oscillations; R = rotational oscillations; n_i' = number of hydroxyl vibrations.

Table III shows the distribution of the several frequencies under the various classes. Group theoretical analysis shows that there should appear in the Raman effect 9 vibrations, *i.e.*, three each under the classes A_g, B_{2g}, and B_{3g}. We have already seen that the symmetric internal vibration of the hydroxyls belongs to class A_g. Since the three pure translations of the unit cell belong to the classes B_{1u}, B_{2u} and B_{3u} and since no rotational oscillations can occur about the two-fold axis $C_2(z)$ owing to the hydroxyl groups lying parallel to it, the remaining two vibrations under A_g should be of the translatory type. These can be described as arising from the translatory movements of the sodiums and hydroxyls along the z-axis and symmetrically with respect to the centre of inversion. Considerations analogous to the above show that there should be one rotational oscillation under each one of the classes B_{2g} and B_{3g}, the rotations of the hydroxyl groups being respectively about the y and x-axes. Also, there occur two oscillations under each one of the classes B_{2g} and B_{3g} due to the translatory movements (symmetrically with respect to the centre of inversion) of the Na's and OH's along the x and y-axes respectively.

Contrary to the theoretically expected number of eight, only five frequency shifts have been observed. However, a tentative explanation for the smaller number of observed shifts can be offered. The unit cell parameters determined by Ernst for NaOH show that for this orthorhombic

structure the lattice constants a and b are equal to each other within the limits of experimental error. Hence, each of the double layers possesses tetragonal symmetry, the four-fold axis (*i.e.*, c-axis) being shifted by $a/2$ as we pass from one double layer to another. If we neglect the weak coupling between the double layers which are far apart, it may easily be seen that the translatory movements along x and y axes (and also the rotational oscillations about the x and y-axes) would be of nearly the same frequency. In other words, the corresponding frequencies coming under the classes B_{2g} and B_{3g} would fall too close to each other to be resolved and only five frequencies would be observed.

The assignments of the individual frequency shifts to specific symmetry modes are arrived at from the following considerations. Even though each Na is surrounded by five hydroxyls, it is the hydroxyl nearest to it (along the z-axis) which can be identified as forming the NaOH group, the other group in the unit cell being related to it by the centre of inversion. (For the sake of convenience we shall describe here only the movements of the atoms in one group, the movements in the other being such that the vibration is symmetric with respect to the centre of inversion.) The translatory oscillations of Na and OH ions can actually be divided into (i) the movements of the Na and the hydroxyl against each other and (ii) the movement of the NaOH group as a whole. The former would be of high frequencies and amongst them the movement of the Na against the OH along the z-axis would be of highest frequency and hence the shift at 380 cm.$^{-1}$ is identified with this mode (A_g). Similar translatory oscillations of the Na's and OH's of the unit cell with their displacements parallel to the x and y axes would be of lower frequency and the frequency shift at 291 cm.$^{-1}$ therefore represents these modes (B_{2g} and B_{3g}). The translatory movements of the NaOH group as a whole parallel to the x, y and z axes would be of low frequencies and are to be identified with the shifts 98 and 114 cm.$^{-1}$ Amongst these, the more intense line at 98 cm.$^{-1}$ conceivably arises due to the overlap of the modes B_{2g} and B_{3g} involving movements parallel to x and y axes and the other shift at 114 cm.$^{-1}$ due to movements normal to the layer (A_g). The rather strong shift at 215 cm.$^{-1}$ is identified with the rotational oscillations about the y and x-axes (*i.e.*, classes B_{2g} and B_{3g}) since rotational oscillations of anisotropic groups usually give rise to strong Raman shifts. The appearance of the translatory oscillations with adequate intensity is due to the coupling generally present between translatory and rotatory symmetry modes.

Sodium hydroxide undergoes a structural transformation at about 300° C. The crystal structure of the high temperature form is not known. Whether the exceptionally large widths of the lines 215 cm.$^{-1}$ and 380 cm.$^{-1}$

are connected with this transformation (as in the case of the 206 cm.$^{-1}$ Raman line in α-quartz) is not quite clear.

(iii) *The monohydrates of sodium and potassium hydroxides.*—As a result of the numerous groups present in the unit cells of these two crystals, the number of lattice vibrations theoretically Raman active in the case of $NaOH.H_2O$ and $KOH.H_2O$ is quite large, being respectively 56 and 28. Nine lattice frequencies have been observed in $NaOH \cdot H_2O$ and eight in $KOH.H_2O$. The following are plausible reasons for the small number of shifts observed. (i) Several of the allowed lines may be inherently weak. (ii) The large separation between the groups as in the case of the layered structure of $NaOH \cdot H_2O$ indicates that the number of actual units that are responsible for the observed frequencies might be fewer than that rigorously considered. (iii) As is evident from the fact of the low transition points of the hydrates, the water molecules are feebly bound to the structure and their translatory and rotatory movements would be of very low frequencies and weak intensities. From a comparison of the spectrum of $NaOH.H_2O$ with that of NaOH it may be seen that there exists a correspondence between some of the frequencies. In particular, the intensities of the shifts 240 cm.$^{-1}$ in $NaOH \cdot H_2O$ and 220, 232 cm.$^{-1}$ in $KOH \cdot H_2O$ and their correspondence to the strong 215 cm.$^{-1}$ shift in NaOH indicate that they also arise most probably due to the rotational oscillations of the hydroxyls. The shifts at 213 and 243 cm.$^{-1}$ observed in $LiOH \cdot H_2O$ may also be due to the rotational oscillations of the hydroxyls.

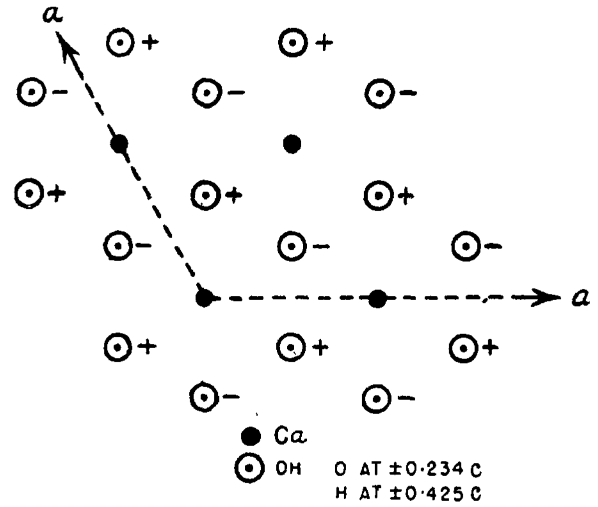
(iv) *Calcium hydroxide.*—In view of the fact that $Ca(OH)_2$ figures in our discussions in Part III also, the group characters and selection rules of the irreducible representations of the group D_{3d} to which it belongs are reproduced in Table IV. In it are also shown the number of vibrations under the different classes and their resolution into the translatory, rotatory and internal oscillations. Figure 2 in the text illustrates the structure of $Ca(OH)_2$.

Group theoretical analysis shows that the Raman spectrum should exhibit four frequency shifts, two each under the classes A_{1g} and E_g. We have already identified the internal vibrations of the hydroxyls symmetrically with respect to the centre of inversion as belonging to A_{1g}. The remaining frequency under A_{1g} represents the translatory movements, along the c-axis, of the hydroxyls symmetrically with respect to the centre of inversion. Similar symmetric translational oscillations of the hydroxyls in the plane perpendicular to the c-axis fall under the doubly degenerate class E_g. The rotational oscillations of the hydroxyls about any axis lying in the plane

TABLE IV

D_{3d}	E	$2C_3$	$3C_2$	i	$2S_6$	$3\sigma_d$	n_i	T	T′	R	n_i'	Activity
A_{1g} ..	1	1	1	1	1	1	2	0	1	0	1	R.E. only
A_{2g} ..	1	1	−1	1	1	−1	0	0	0	0	0	Inactive in both
E_g ..	2	−1	0	2	−1	0	2	0	1	1	0	R.E. only
A_{1u} ..	1	1	1	−1	−1	−1	0	0	0	0	0	i.a. both
A_{2u} ..	1	1	−1	−1	−1	1	3	1	1	0	1	I.R. only
E_u ..	2	−1	0	−2	1	0	3	1	1	1	0	I.R only
U_R ..	5	5	1	1	1	5						

R.E. = Raman effect; I.R. = infra-red; i.a. = inactive.



TEXT-FIG. 2. Projection of the structure of Ca $(OH)_2$ in a plane normal to the c-axis. + and − indicate the levels of the atoms. The hydroxyl groups are parallel to c. The Ca atoms are in the plane passing through the origin.

perpendicular to the c-axis belong to the class E_g. The calciums do not take part in any of the Raman active vibrations owing to their being located at the centres of inversion.

As indicated by theory, in the case of Ca $(OH)_2$ three lattice frequency shifts are observed at 247, 282 and 359 cm.$^{-1}$ The observed shifts, however, fall within too limited a range of frequencies to permit their identification with specific symmetry modes.

(v) *Octahydrated hydroxides of strontium and barium.*—Since the lattice spectra of Sr $(OH)_2 \cdot 8\ H_2O$ and Ba $(OH)_2 \cdot 8\ H_2O$ exhibit respectively only eight and nine frequency shifts (which are far fewer than would be indicated by the presence of the large number of groups in the unit cell) and also since the positions of the hydrogen atoms have not been determined, no attempt is made here to enumerate the frequencies theoretically. The reasons advanced in the cases of $NaOH \cdot H_2O$ and $KOH \cdot H_2O$ for the meagreness of the results observed in this region of the spectrum apply equally well to these cases also.

The unit cell parameters of Ba $(OH)_2 \cdot 8\ H_2O$ are given by $a = 9 \cdot 35$; $b = 9 \cdot 28$; $c = 11 \cdot 87$ A.U., and $\beta = 99°$. The substance exhibits a small optic axial angle of 8° 33′. The unit cell dimensions of Sr $(OH)_2 \cdot 8\ H_2O$ are $a = 9 \cdot 00$; $c = 11 \cdot 58$ A.U. and are comparable to that of Ba $(OH)_2 \cdot 8\ H_2O$. It appears highly likely that the structure of Ba $(OH)_2 \cdot 8H_2O$ is only a slightly distorted version of Sr $(OH)_2 \cdot 8\ H_2O$ and hence a comparison of their spectra is of interest.

In the lattice frequency region and in the high frequency region as well, a fair agreement between the frequency shifts of Ba $(OH)_2 \cdot 8\ H_2O$ and Sr $(OH)_2 \cdot 8\ H_2O$ may be noticed to exist. An explanation for the correspondence between the two spectra is possible since the unit cell parameters in the two cases are comparable and also since the monoclinic group C_{2h} to which Ba $(OH)_2 \cdot 8\ H_2O$ belongs is a subgroup of the tetragonal group D_{4h} of Sr $(OH)_2 \cdot 8\ H_2O$. From the point of view of Raman effect studies, only the following correlations existing between the five symmetric representations of D_{4h} and the two symmetric representations of C_{2h} are of interest. They are: (1) $A_{1g} \rightarrow A_g$; (2) $A_{2g} \rightarrow B_g$; (3) $B_{1g} \rightarrow A_g$; (4) $B_{2g} \rightarrow B_g$; (5) $E_g \rightarrow A_g + B_g$. In the group D_{4h} only the four symmetric classes A_{1g}, B_{1g}, B_{2g} and E_g are Raman active, A_{2g} being inactive. In C_{2h} both A_g and B_g are Raman active. Hence, as a result of the lowering of the symmetry from D_{4h} to C_{2h} the inactive vibrations corresponding to A_{2g} of D_{4h} become active and the doubly degenerate vibrations (E_g) split into two. The lines under the totally symmetric class are often observed to be of considerable intensity. The transformation of B_{1g} of D_{4h} into A_g of C_{2h} might hence mean that the corresponding line observed with the substance of lower symmetry would appear with enhanced intensity. The appearance of the hydroxyl frequency at 3529 cm.$^{-1}$ rather strongly in Ba $(OH)_2 \cdot 8\ H_2O$ in contrast to its weak counterpart at 3536 cm.$^{-1}$ in Sr $(OH)_2 \cdot 8\ H_2O$ is therefore intelligible. The appearance of doublets in the lattice spectrum of Ba $(OH)_2 \cdot 8\ H_2O$ corresponding to single frequencies observed in Sr $(OH)_2 \cdot 8\ H_2O$ (*e.g.*, 230 → 205–220; 297 → 276 to 287) is obviously due to the splitting arising from the

removal of degeneracy. Here again, the two frequencies 205 and 220 cm.$^{-1}$ (corresponding to the line 215 cm.$^{-1}$ due to rotational oscillations of the hydroxyls in NaOH) represent the rotational oscillations of the hydroxyls, the former which is more intense belonging to the class A_g and the latter to B_g. The frequency shifts 230 and 297 cm.$^{-1}$ belong to the doubly degenerate class E_g of D_{4h}. It is evident that these corresponding frequencies represent predominantly the translatory and rotatory oscillations of the hydroxyls since the vibrations in which the bariums and strontiums participate would be of different frequencies owing to the large difference in their atomic masses.

The author's grateful thanks are due to Professor Sir C. V. Raman, F.R.S., N.L., for the valuable discussions he had with him.

5. Summary

The paper discusses the results reported in Part I in relation to the crystal structures of the substances. From a comparison of the facts with the theoretically expected picture of the spectrum, the salient features exhibited by the spectra have been explained, in addition to identification and assignment of the frequencies wherever feasible. The progressive variations in the hydroxyl shifts exhibited by the different hydroxides and their hydrates are explicable as due to the interactions of the hydroxyls with each other and also with their neighbouring water molecules in the case of the hydrates.

6. References

1. Ernst, T. .. *Zeit. Physik. Chem.*, 1933, **20 B**, 65.
2. Pepinsky, R. .. *Phys. Rev.*, 1939, **55**, 1115. (*See* also *Strukturbericht*, Band VII, 1939, pp. 4 *et seq.*).
3. Ernst, T. .. *Angew. Chem.*, 1948, **60 A**, 77.
4. Busing, W. R. .. *J. Chem. Phys.*, 1955, **23**, 933.
5. Wunderlich, J. A. .. *Bull. Soc. Fran. Miner., Crist.*, 1958, **81**, 287.
6. Clastre, J. .. *Acta Cryst.*, 1954, **7**, 638.
7. Bragg, W. L. .. *Atomic Structure of Minerals*, Oxford University Press (1937), pp. 108.
8. Busing, W. R. and Levy, H. A. *J. Chem. Phys.*, 1957, **26**, 563.
9. Petch, H. E. .. *Canad. Jour. Phys.*, 1957, **35**, 983.
10. Smith, H. G. .. *Acta Cryst.*, 1953, **6**, 604.
11. Ramaseshan, S. .. Private communication.
12. Megaw, H. D. .. *Zeit. Krist.*, 1934, **87**, 185.
13. Bhagavantam, S. .. *Proc. Ind. Acad. Sci.*, 1941, **13 A**, 543.

14. Levy, H. A. and Peterson, S. W. *Rev. Mod. Phys.*, 1958, **30**, 101.

15. Bernal, J. D. and Megaw, H. D. *Proc. Roy. Soc.*, 1935, **151 A**, 384.

16. Partington, J. R. .. *General and Inorganic Chemistry*, McMillan & Co., London, 1958, pp. 322.

17. Krishnamurti, D. .. *Proc. Ind. Acad. Sci.*, 1959, **50 A**, 223.

18. Wilson, E. B., *et al.* .. *Molecular Vibrations*, McGraw-Hill Book Co., 1955.

19. Herzberg, G. .. *Infra-red and Raman Spectra of Polyatomic Molecules*, D. Van Nostrand Co., Inc., 1945, pp. 280 *et seq.*

20. Bhagavantam, S. and Venkatarayudu, T. *Theory of Groups and Its Application to Physical Problems*, Andhra University, Waltair, 1948.

THE RAMAN AND INFRA-RED SPECTRA OF SOME SOLID HYDROXIDES

Part III. Discussion of the Infra-red Data

By D. Krishnamurti, F.A.Sc.

(*Memoir No.* 119 *from the Raman Research Institute, Bangalore*-6)

Received August 27, 1959

1. Introduction

Comparatively few investigations on the infra-red spectra of the hydroxides appear to have been made and even these mainly concern themselves with the 3 μ region. In the present part we shall discuss the data available for LiOH, $LiOH \cdot H_2O$, NaOH, $Ca(OH)_2$ and $Mg(OH)_2$. Duval and Lecomte (1941) had reported the infra-red spectra of LiOH, $Ca(OH)_2$, $Mg(OH)_2$ and $Al(OH)_3$ for the region 600–1600 cm.$^{-1}$ which falls outside the range of even the first order combinations of most of the lattice frequencies of these substances. In many of the cases studied by them there appear strong maxima in the region of 880 cm.$^{-1}$ and 1440 cm.$^{-1}$ corresponding to the strong absorption maxima usually observed with the carbonates. In view of the meagreness of the data for the lattice frequencies active in the infra-red and also since it appears to us that the substances used by Duval and Lecomte in their study have been contaminated by the atmospheric carbon dioxide to form the carbonates, we do not propose to discuss their data.

2. The Infra-red Spectra of LiOH, $LiOH \cdot H_2O$ and NaOH

(i) *LiOH and* $LiOH.H_2O$.—The infra-red absorption spectra of LiOH and $LiOH.H_2O$ have been recorded by Jones (1954) in the 3 μ region using thin films of the substances, the frequencies of the absorption maxima of the OH vibrations being respectively 3678 and 3574 cm.$^{-1}$ The internal vibrations of the hydroxyls antisymmetrically with respect to the centre of inversion in these two cases belong respectively to the classes A_{2u} of D_{4h} and B_u of C_{2h} and are in each case the only infra-red active internal mode theoretically possible and hence the respective maxima at 3678 and 3574 cm.$^{-1}$ are to be identified with them. That the infra-red frequencies in these cases are only slightly different from the Raman active frequencies is indicative of weak coupling between the hydroxyls. The diminution of the hydroxyl frequency in $LiOH \cdot H_2O$ by about 100 cm.$^{-1}$ is hence largely to be attributed

to the proximity of the water molecules to the hydroxyls. Two other maxima at 3000 cm.$^{-1}$ and 1586 cm.$^{-1}$ respectively due to the stretching and bending vibrations of the water molecules in $LiOH \cdot H_2O$ have also been reported by him, the diminution in their frequencies being due to causes already discussed.

An absorption maximum at 7195 cm.$^{-1}$ has also been observed by Jones in LiOH and identified as the overtone of 3678 cm.$^{-1}$ According to theory this overtone is forbidden in the infra-red and only the combination between the Raman active (3664 cm.$^{-1}$) and the infra-red active (3678 cm.$^{-1}$) vibrations is allowed in the infra-red. Anharmonicity can be expected to be the cause of the violation of the selection rules. The large difference (161 cm.$^{-1}$) between the anticipated value for the overtone and the observed value indicates the presence of considerable mechanical anharmonicity. However, further studies are required to elucidate this feature more fully.

(ii) *Sodium hydroxide.*—The infra-red absorption spectrum of NaOH in the form of thin films has been studied by Busing (1955) and a sharp maximum at 3637 cm.$^{-1}$ is reported by him. This is to be identified with the theoretically expected infra-red active internal vibration of the hydroxyls antisymmetrically with respect to the centre of inversion and belonging to the class B_{1u} of D_{2h}. The small difference of 5 cm.$^{-1}$ between the Raman active and infra-red active hydroxyl frequencies indicates that the coupling between the hydroxyls is very weak.

3. THE INFRA-RED SPECTRA OF $Ca(OH)_2$ AND $Mg(OH)_2$

Unlike the above cases, the infra-red spectra of $Ca(OH)_2$ and $Mg(OH)_2$ exhibit complex structures in the 3 μ region and it was therefore suspected that the actual unit cell of the structure should be larger than that derived from X-ray studies (Mara and Sutherland, 1953). However, reinvestigations of the structures by both X-ray and neutron diffraction methods have only revealed the correctness of the structure proposed earlier. So far, only tentative explanations have been offered for the complex features of the spectra.

It has already been shown in Table IV of Part II that in the cases of $Ca(OH)_2$ and $Mg(OH)_2$, there are possible four infra-red active fundamentals of which one is due to the internal vibrations of the hydroxyls antisymmetrically with respect to the centre of inversion and three others are due to translatory and rotatory oscillations. Besides these, the selection rules for combinations and overtones indicate that the combinations between the Raman active fundamentals and the infra-red active fundamentals (*viz.*, $A_{1g} \times A_{2u}$;

$A_{1g} \times E_u$; $E_g \times A_{2u}$; $E_g \times E_u$) are allowed in the infra-red. The first overtones and first order combinations of the infra-red active frequencies between themselves and similarly the combinations and overtones of the Raman active frequencies are forbidden in infra-red absorption. An absorption maximum can be identified as a combination of the differential type only if the corresponding summational is also observed. The complex features exhibited by the spectra of $Ca(OH)_2$ and $Mg(OH)_2$ are explicable on the basis of these criteria.

(i) *Calcium hydroxide.*—The infra-red absorption spectrum of $Ca(OH)_2$ has been reported by Busing and Morgan (1957). Table I shows the frequencies of the absorption maxima and the nature of their origin. The

TABLE I

Absorption maxima in $Ca(OH)_2$ and their interpretation

Absorption maximum in cm.$^{-1}$	Their suggested origin	Calculated value in cm.$^{-1}$
4164	3616+282+263	4161
3932	3644+282; 3616+322	3926; 3938
3899	3644+247	3891
3877	3616+263	3879
3720	3616+98	3714
3644	A_{2u}	..
3629*	?	..
3523	3616−98	3518
3364	3644−282	3362
3333†	3616−263	3353
3294	3616−322	3294
3073	3616−282−263	3071

* Very weak; † See text for alternative explanation.

prominent maximum at 3644 cm.$^{-1}$ observed by them, when the crystal is orientated such that the incident beam is inclined to the c-axis, represents the hydroxyl stretching vibrations (antisymmetrically with respect to the centre of inversion) under the class A_{2u}. The numerous other maxima are explicable as arising due to (i) the combinations of the frequency 3644 cm.$^{-1}$ with the Raman active lattice frequencies, and (ii) combinations of the Raman active frequency 3616 cm.$^{-1}$ with the infra-red active lattice frequencies. The actual interpretation of the facts was arrived at by considering the combinations of the former type first and then by choosing from amongst the remaining maxima those pairs of frequencies whose mean value was near 3616 cm.$^{-1}$ By this process it was possible to discover also the three lattice frequencies (98, 263 and 322 cm.$^{-1}$) active in the infra-red. The calculated and observed values shown in Table I are in fair agreement except for the frequency at 3333 cm.$^{-1}$ The only other explanation that could be given for this maximum, without disturbing the consistent scheme of assignments arrived at for the other frequencies, is that it is a combination, forbidden by selection rules, *viz.*, 3616 − 282; the corresponding summational (3616 + 282) is also observed at 3899 cm.$^{-1}$ The infra-red activities of these are presumably to be attributed to their contiguity to strongly active infra-red frequencies. The appearance of the Raman frequency 282 cm.$^{-1}$ in combinations (differential and summational) with the infra-red active frequency of 3644 cm.$^{-1}$ vindicates our earlier identification of it as a genuine Raman shift of $Ca(OH)_2$.

(ii) *Magnesium hydroxide*.—The infra-red absorption spectrum of $Mg(OH)_2$ has been investigated in detail recently by Mara and Sutherland (1953) and also by Benesi (1959). Using fine powders Benesi has observed a single sharp intense absorption maximum at 3698 ± 2 cm.$^{-1}$ corresponding to the internal vibrations of the hydroxyls (antisymmetrically with respect to the centre of inversion) belonging to the class A_{2u}. Mara and Sutherland who used single crystals of $Mg(OH)_2$ report numerous maxima in addition to a maximum at 2·71 μ (corresponding to 3698 cm.$^{-1}$) which appears when the c-axis of the crystal is inclined to the incident beam. The explanation of the several maxima shown in Table II is rather indirect since neither the Raman active nor the infra-red active lattice frequencies have so far been observed. The numerous maxima arise evidently due to combinations, (i) between the Raman active frequency 3651 and three infra-red active lattice frequencies and (ii) between the infra-red active frequency 3698 cm.$^{-1}$ and the three Raman active lattice frequencies. These combinations should appear on either side of the frequency region 3651–3698, the mean value of the corresponding summational and differential frequencies being centred

TABLE II

Absorption maxima in $Mg(OH)_2$ and their interpretation

Observed maxima		Their suggested origin	Calculated values	
in μ	in cm.$^{-1}$		in cm.$^{-1}$	in μ
2·06	4854	Combinations of higher order	..	..
2·19	4566	do.	..	..
2·24	4464	$3698+2\times 380$	4458	2·243
2·32	4310	$3651+380+293$	4324	2·313
2·45	4082	$3698+380$	4078	2·452
2·47	4049	$3651+390$	4041	2·474
2·49	4016	$3698+324$	4022	2·486
2·53	3953	$3651+293$	3944	2·535
2·64	3788	$3651+127$	3778	2·647
2·71	3690	A_{2u}	..	..
2·74	3650	A_{1g} ?	..	..
2·83	3534	$3651-127$	3524	2·838
2·97	3367	$3651-293$; $3698-324$	3358; 3374	2·978; 2·964
3·01	3322	$3698-380$	3318	3·014
3·06	3268	$3651-390$	3261	3·066
3·08	3247*	..	..	..

* See text for explanation.

about either 3651 or 3698 cm.$^{-1}$ The identification of the corresponding summational and differential frequencies is facilitated by the similarity of their behaviour in the spectra obtained for two different orientations of the crystal. For example, the maxima at 2·45 μ and 3·01 μ are observed only when the incident beam is perpendicular to the *c*-axis and not when it is parallel to the *c*-axis.

The wavelengths of the absorption maxima are given by Mara and Sutherland to only two places of decimals and it can be seen that the frequency 3690 observed by them differs (by 8 cm.$^{-1}$) from the more accurate measurement of Benesi. In view of this possibility of errors of the order of 10 cm.$^{-1}$ in the observed values, the agreement between the calculated and observed values of Table II should be considered quite satisfactory. The analysis of data reveals that there should be observed in the case of $Mg(OH)_2$ the following lattice frequencies: (*a*) 324 and 380 cm.$^{-1}$ in the Raman effect; (*b*) 127, 293 and 390 cm.$^{-1}$ in the infra-red absorption. The value of the third Raman active lattice frequency might probably be 451 cm.$^{-1}$ as shall be seen below. It is possible to explain the maximum observed at 3247 cm.$^{-1}$ as a differential, *i.e.*, 3698—451, the frequency 451 cm.$^{-1}$ being assumed to be of the Raman active type. In that case, the corresponding summational also should be observed at 4149 cm.$^{-1}$, *i.e.*, at 2·41 μ. But, such a maximum has not been observed, probably because of its falling close to the strong absorption at 2·45 μ and being obscured by it. However, the above is only a tentative explanation. It is rather surprising that there should be observed at 3650 cm.$^{-1}$ in the infra-red, a frequency closely agreeing with the Raman active frequency, in violation of the rule of mutual exclusion.

The crystal structure of $Mg(OH)_2$ is a more closely packed one than that of $Ca(OH)_2$ as is evident from a comparison of their lattice constants, which are given by:

(i) $Mg(OH)_2$: $a = 3 \cdot 12$; $c = 4 \cdot 73$ A.U.

(ii) $Ca(OH)_2$: $a = 3 \cdot 59$; $c = 4 \cdot 91$ A.U.

Consequently, the interionic forces in $Mg(OH)_2$ would be stronger than in $Ca(OH)_2$ and the fact that the lattice frequencies of $Mg(OH)_2$ are higher than those of $Ca(OH)_2$ is therefore not surprising. The lower atomic mass of Mg would also be responsible for higher frequencies, in the case of modes involving movements of the magnesiums.

It is interesting to note that in all the cases discussed herein, the infra-red active hydroxyl frequencies are always higher than the Raman active hydroxyl frequencies. In the latter case, the dipole moments produced at the two hydroxyls related by the centre of inversion are opposite to each other resulting in their cancellation and hence the vibration is infra-red inactive. On the other hand, in the case of the infra-red active vibrations, the dipole moments produced at the two hydroxyls are additive. Because of the fact that there is a displacement of charge, or in other words, a polarization of the medium taking place in the case of the infra-red active vibration, the potential energy and hence the restoring force for this mode is

greater than in the case of the Raman active mode. The higher value of the infra-red active hydroxyl frequency is hence intelligible.

In conclusion, the author wishes to express his deep sense of indebtedness to Professor Sir C. V. Raman, F.R.S., N.L., for the encouragement and valuable suggestions that he gave during the course of this investigation.

4. Summary

In each of the cases of LiOH, $LiOH \cdot H_2O$, NaOH, $Ca(OH)_2$ and $Mg(OH)_2$ the stretching vibrations of the hydroxyls antisymmetrically with respect to the centre of inversion give rise in infra-red absorption to a single sharp intense maximum at about 3650 cm.$^{-1}$ as is to be theoretically expected. In all cases, the infra-red active hydroxyl frequency is found to be higher than the corresponding Raman active mode and the origin of this feature is explained.

The spectrum of $Ca(OH)_2$ recorded with single crystals exhibits on either side of the principal maximum due to the hydroxyl vibration, numerous other maxima which have hitherto remained unexplained. Some of them arise due to combinations between the observed Raman active lattice frequencies and the infra-red active hydroxyl frequency. The rest are explicable as combinations between the Raman active hydroxyl frequency and infra-red active lattice frequencies. Features similar to those observed in the spectrum of $Ca(OH)_2$ are exhibited by $Mg(OH)_2$ also, and have been explained in an analogous fashion. No experimental data regarding the infra-red active lattice frequencies of $Ca(OH)_2$ and the Raman active and infra-red active lattice frequencies of $Mg(OH)_2$ are available. The values of these are also revealed by the above analysis.

5. References

Benesi, H. A.	.. *J. Chem. Phys.*, 1959, **30**, 852.
Busing, W. R.	.. *Ibid.*, 1955, **23**, 933.
——— and Morgan, H. W.	.. *Ibid.*, 1958, **28**, 998.
Duval and Lecomte	.. *Bull. Soc. Chim., France (V)*, 1941, **8**, 713.
Jones, L. H.	.. *J. Chem. Phys.*, 1954, **22**, 217.
Krishnamurti, D.	.. *Proc. Ind. Acad. Sci.*, 1959, **50 A**, 232.
Mara, R. T. and Sutherland, G· B. B. M.	*J. Opt. Soc. Am.*, 1953, **43**, 1100.

Reprinted from "The Proceedings of the Indian Academy of Sciences",
Vol. L, No. 6, Sec. A, 1959

X-RAY STUDY OF THE STRUCTURE OF MOONSTONES

BY A. JAYARAMAN, F.A.SC.
(*Memoir No.* 120 *from the Raman Research Institute, Bangalore*-6)

Received October 24, 1959

1. INTRODUCTION

MOONSTONES which represent the potash-soda series of mixed feldspars lying in the compositional range Or_{85} Ab_{15} to Or_{40} Ab_{60}, exhibit a characteristic optical effect, commonly known as schiller. To all external appearances, they are monocrystals with symmetry elements characteristic of the monoclinic system. The optical phenomenon displayed by them, however, betray their complex nature, for the phenomenon arises due to optical heterogeneity of the medium. X-ray methods may therefore be expected to throw light with regard to the precise nature of the material responsible for the heterogeneity. Indeed, earlier X-ray investigations of Kozu and Endo (1921), Chao, Smare and Taylor (1939), Chao and Taylor (1940) and Ito (1950) have shown that moonstones possess a complex structure; the soda-feldspar having segregated as a distinct crystalline phase with triclinic symmetry, from the monoclinic potash-feldspar. The fact that the existence of the former feldspar as a separate phase is directly connected with the schiller phenomenon is shown by heat-treatment studies on moonstones. In recent years, X-ray studies by Mackenzie and Mackenzie and Smith (1955, 1956, 1958) on alkali feldspars have shown that there exist high and low temperature perthites. Further, the unmixed soda-phase has a considerable amount of potash feldspar in solid solution as is found from a^* and γ^* values of the former phase. All previous X-ray studies have established the fact that while the potash feldspar phase exist with monoclinic symmetry, the soda-feldspar crystallites in the moonstones have triclinic symmetry and are twinned either according to pericline or albite law. However, the explanations advanced with regard to the factor influencing the type of twinning assumed by the soda-feldspar crystallites are contradictory and have not provided the correct answer to this interesting question. Further, there are some points with regard to the structure of Korean Moonstones needing clarification, in view of the different interpretations given by Chao and Taylor on the one hand and by Ito on the other.

The optical effect exhibited by moonstones was investigated in great detail in this laboratory several years ago and an account of the same appeared

in these *Proceedings* (1950, 1953). The need for an independent X-ray investigation was keenly felt at that time to elucidate on the complex nature of moonstones whose optical behaviour was thoroughly understood. For various reasons such a study could not be carried out then. Recently another type of moonstone from the Coimbatore District in South India came to our notice which exhibits, apart from a golden yellow schiller, other highly characteristic optical effects, namely, polarised transmission and forward diffusion effects. Subsequent to our previous studies, a fair number of Korean moonstones exhibiting schiller were added to our collection. These circumstances and for the reasons already set forth, a thorough investigation by X-ray methods was undertaken recently on all the above-mentioned materials by single crystal methods. Of these, the Coimbatore moonstone is studied for the first time. As a precise investigation was called for, the Weissenberg method was extensively employed along with rotation methods. Both natural and heat-treated materials were studied and accurate measurements of the lattice constants were obtained. The results have greatly aided in putting forth an explanation for the differences in the twinning behaviour, based on the differences between the lattice constants of the constituent feldspar in moonstones. Further, the studies by the Weissenberg method have led to a determination of the precise nature of the crystallographic relationships of the two feldspars and the structural characters of the soda-feldspar phase. In this communication the results of the experimental study and the conclusions drawn from them would be set forth in detail.

2. Rotation Diagrams of Moonstones

In the oscillation and rotation photographs of all moonstones, two kinds of reflections are distinguishable; a set of strong, sharp reflections and a set of weak and somewhat diffuse ones. They evidently have different origin, the stronger ones arising due to the potash-feldspar phase and the latter due to soda-feldspar. Certain common features are noticeable in the X-ray rotation photographs of the three moonstones (Coimbatore, Ceylon and Korean varieties). The principal reflections in the *b*-axis rotation photographs are symmetrically disposed with respect to the zero layer, revealing thereby the presence of a symmetry plane perpendicular to that axis. The *a*-axis photographs of all moonstones examined consist of a double set of layer lines, one set comprising of strong reflections and another of weaker reflections. The latter are spaced wider than the former and the corresponding axes lie very nearly parallel. In every case, the morphological axes of the moonstone are found to coincide with the crystallographic axes of the potash feldspar phase.

The *b*- and *c*-axis rotation photographs divide these three moonstones into two types. In one type, while the X-ray photograph for the *c*-axis rotation exhibits features characteristic of a simple layer line diagram, the *b*-axis rotation photograph is complex. In the latter, strong reflections fall accurately upon layer lines and weak subsidiary reflections accompany these. In general, each principal reflection is accompanied by a pair of satellites, one lying above and the other below the layer line, on curves of constant ζ. The distribution of the subsidiary reflections is characteristic of an oscillation photograph with misset axes and their symmetrical appearance about the layer line indicates the presence of two such axes, equally inclined to the principal axis. In the second type, the position is reversed, namely, that while the *b*-axis rotation picture has the characteristics of a simple layer line diagram, the rotation diagram for the *c*-axis is complex. The distribution of the subsidiary reflections in the photograph is indicative of the existence of two subsidiary *c*-axes, symmetrically inclined to the principal axis.

To the first type belong the Ceylon and Coimbatore moonstones and to the second type the Korean variety. Figures 1 and 2 reproduce *b*-axis oscillation photographs of the Coimbatore and Korean moonstones respectively. In recording these, the crystal was oscillated through a range of 15° with (001) plane set parallel to the X-ray beam at the middle of the oscillation range (*see* Mackenzie and Smith, 1955). In the oscillation photograph of the Korean moonstone, the principal reflection can be seen accompanied by weaker subsidiary reflections. The features described for the Coimbatore and Ceylon moonstones are clearly to be seen in Fig. 1.

The foregoing facts become intelligible, if it is assumed that in all moonstones, the potash feldspar exists as a truly monoclinic component and that the soda-feldspar crystallites possess triclinic symmetry and are arranged in the mutual orientation characteristic of albite twin law in the case of Coimbatore and Ceylon moonstones, and in the orientation characteristic of the pericline twin law in the case of Korean moonstone.

3. Weissenberg X-Ray Diagrams

Zero layer Weissenberg photographs about *a*, *b* and *c* axes were recorded for all the varieties of moonstones, employing CuK_α radiation and a Unicam Weissenberg X-ray Goniometer.

a-axis.—In the *a*-axis photograph of the Coimbatore moonstone, two subsidiary reflections appear close to the principal ones from (*okl*) and (*ool*) planes, while in the case of Korean moonstones, the *oko* and *okl* spots are accompanied by two such reflections. Single spots appear for the *oko*

reflections in the former moonstone and for the *ool* reflections in the latter. The *a*-axis zero layer photograph of the two moonstones are reproduced in Figs. 3 and 4. In the photograph for the Korean moonstone the *oko* reflections apear as long streaks but these actually consist of three spots merging into each other. The triple character of the reflections can be seen in *okl* reflections.

b-axis.—In photographs for the Coimbatore moonstones, the principal reflections from the (*ool*) and (*hol*) planes have two close satellites and the corresponding reflections in the case of Korean moonstone consist of single and double spots. In both cases the *hoo* reflections are accompanied by single subsidiary reflections which lie more or less in the same line as the principal reflection.

c-axis.—In the *c*-axis zero layer photograph of the Coimbatore moonstone all the reflections except those from (*oko*) planes are doubled. But in the photographs for the Korean moonstone it is seen that the principal reflections from the (*hko*) and (*oko*) planes have two subsidiary reflections, the *hoo* reflections appearing doubled.

The X-ray photographs of Ceylon moonstones of both the blue and white schiller variety exhibit characters similar to that described for the Coimbatore variety. The differences in the features observable on the Weissenberg photographs arise due to the different twinning laws of the soda-feldspar phase. The two individuals of the albite twin have the (010) plane common and are related to each other by a rotation through 180° about the normal to this plane. While this operation leaves the b^* axes coincident, the c^* axes of the reciprocal lattice do not coincide and are symmetrically inclined with respect to the c^* axis of the monoclinic phase. Therefore planes involving l indices of the soda-feldspar have two points in the reciprocal lattice, giving rise to two subsidiary reflections. In the case of pericline twinning, the individuals are related by a rotation through 180° about the *b*-axis, which operation, while bringing the a^* and c^* axes into coincidence leave the b^* axes non-coincident. The latter are symmetrically inclined with respect to the b^* axis of the monoclinic phase. Now, the planes involving k indices are doubled and give rise to two subsidiary reflections. The (*hoo*) spacings differ for the two feldspars and therefore reflections from these are doubled.

A noteworthy feature observable on Weissenberg photographs is that only single spots appear from the triclinic soda-feldspar phase accompanying the principal reflections from (*hoo*) planes. This result leads to the important conclusion that the interaxial angle γ is very nearly equal to 90° for the soda-feldspar phase in these moonstones.

The zero and *n* layer Weissenberg photographs of the Korean moonstone rotated about the *b*-axis reveal that a monoclinic potash phase and a pericline twinned soda-phase coexist in it. No evidence for the presence either of a triclinic potash phase or monoclinic soda-phase could be found in the photographs. This is in agreement with the results obtained by Chao and Taylor (*loc. cit.*) for Korean moonstone.

4. X-Ray Patterns of Heat-treated Moonstones

When moonstones are heated to about 850° for a few hours, the schiller vanishes. This fact was reported earlier by E. Spencer (*loc. cit*). Some of the heat-treated specimens were examined by Taylor and collaborators (*loc. cit.*).

In the present series of studies, the moonstones were heat-treated, and their X-ray patterns were recorded both by rotation and Weissenberg methods. The heat-treatment was carried out at 900° C. for about three hours and the specimens were then allowed to cool to room temperature. In X-ray photographs of heat-treated moonstones, all subsidiary reflections characteristic of the untreated material vanish. The detailed characteristics of the rotation diagrams which consist of single set of layer lines, and the Weissenberg photographs reveal that the heat treated material possess monoclinic symmetry. The lattice constants of the heat-treated moonstones were obtained from a measurement of the high-order reflections in the respective Weissenberg photographs and are presented in Table I along with the data for the potash-feldspar phase of the untreated material.

Table I

Lattice Constants of Natural and Heat-treated Moonstones in A.U.

Lattis Constant	Coimbatore moonstone		Ceylon moonstone		Korean moonstone	
	Natural	Heat-treated	Natural	Heat-treated	Natural	Heat-treated
a^*	7·70	7·58	7·72	7·619	7·63	7·50
b^*	12·94	12·94	12·84	12·84	12·94	12·94
c^*	6·40	6·42	6·42	6·45	6·39	6·40
a	8·57	8·46	8·59	8·47	8·49	8·34
b	12·94	12·94	12·84	12·84	12·94	12·96
c	7·20	7·14	7·17	7·17	7·12	7·13

It will be seen from a scrutiny of the data presented in the table that a significant contraction along the *a*-axial dimension has taken place on heat-treatment. This change is found to be proportional to the amount of soda-feldspar that has gone into solution. In estimating the contraction, the value for the natural material should be taken as 8·61 Å which is the value for pure orthoclase. If a certain percentage of soda-feldspar has already gone into solution in the potash phase of the untreated material, the actual value for the latter would be brought down. This is what has happened in the case of Korean moonstone. The difference in the *a*-axial dimension between heat-treated Korean moonstone and orthoclase is nearly double that of such differences for the other two moonstones, as is to be expected, in view of the higher percentage of soda-feldspar present in it, nearly 60% as against 30 to 33% in the others.

5. Discussion of the Results

The results obtained raise the interesting question, namely, what is the factor that is responsible for the soda-feldspar crystallites in the moonstones assuming a orientation characteristic of albite law in one case and pericline law in the other. Chao and Taylor (*loc. cit*) classified Spencer's specimens into two groups; those having albite twinned soda-phase whose lattice angles were close to low-albite, and those having pericline twinned soda-phase whose lattice angles did not agree with low-albite. They were not aware of the existence of high albite at that time which was proved beyond doubt by Tuttle and Bowen (1950), and gave a tentative explanation for this difference based on bulk chemical composition. Ito and Sadanaga (1952) put forward a similar suggestion. Laves (1952) thought that the type of twinning was controlled, depending upon whether the soda-phase has low or high albite structure. This suggestion was however contradicted by Smith and Mackenzie (1954).

Table II presents the *a*-axial dimensions of the potash phase and the lattice angles α^* and γ^* of the soda-feldspar phase for the three moonstones studied. The latter were obtained from measurements on Weissenberg photographs. The lattice angles for low albite are also given in the last column.

Mackenzie and Smith have shown that from the α^* and γ^* values of the soda-phase, the chemical composition of that phase could be determined. As remarked earlier, the present investigation shows that the composition of the monoclinic potash phase is revealed by the *a*-axial dimension of that phase. The *a*-axial dimension of the Korean moonstone whose bulk chemical composition is $Or_{40}\,Ab_{60}$ indicates that the potash phase in it holds an appreciable proportion of soda-feldspar, amounting to nearly 25 to 30% in solid

TABLE II

Phase	Lattice constant	Korean moonstone	Ceylon moonstone	Coimbatore moonstone	Albite
Potash ..	a	8·49Å	8·59Å	8·57Å	..
Soda ..	α^*	88° 30′	86° 30′	87° 24′	86° 24′
Soda ..	γ^*	90°	90° 20′	90° 20′	90° 29′

solution. The α^* and γ^* values of the soda-phase in the same moonstone correspond to a chemical composition $Or_{30}\,Ab_{70}$.

It will be seen from Table II that in the case of Korean moonstones, the differences in respect of both the *a*-axial dimensions and the lattice angles α^* and γ^* between the two constituent feldspars are lessened, this essentially arising out of solid solution of an appreciable proportion of one phase in the other. Such circumstances favour pericline twinning of the soda-phase. On the contrary, these differences are greater in the case of Coimbatore and Ceylon moonstones and the two feldspar phases are very close to orthoclase and low-albite. When the soda-feldspar phase thus demands full-triclinic symmetry of low-albite and the *a*-axial dimension of the potash phase correspond to that of orthoclase, twinning takes after the albite law. The principal factor that seems to influence the nature of twinning is therefore the narrowing down of the differences in respect of the *a*-axial dimension of the two feldspars.

In conclusion the author wishes to express his sincere thanks to Sir C. V. Raman, F.R.S., N.L., for his kind interest in this work.

SUMMARY

X-ray studies on three different types of moonstones, namely, Ceylon, Coimbatore and Korean varieties have been made by single crystal rotation and Weissenberg methods. The results are consistent with the explanation that the triclinic soda-feldspar crystallites which has segregated out of monoclinic potash feldspar are twinned according to albite law, in the former two, while the twinning is according to pericline law in the latter moonstone. The view is put forward that the narrowing down of the differences, firstly of the *a*-axial dimension between the soda-feldspar and potash feldspar phases, and secondly between the lattice angles of the two feldspars result in pericline

twinning of the soda-phase, while albite twinning results when the soda-phase demands full triclinic symmetry and the potash phase has the *a*-axial dimension corresponding to that of orthoclase. The detailed features of the Weissenberg X-ray diagrams of Korean moonstones reveal that there exist in it a monoclinic potash phase and a pericline twinned triclinic soda-phase, in agreement with the results obtained by Chao and Taylor. X-ray results also indicate that the lattice angle γ of the triclinic soda-phase in these moonstones is very nearly equal to 90°.

REFERENCES

1. Kozu, S. and Endo, Y. .. *Sci. Rept. Tohoku Univ. Ser.* 3, 1921, **1,** 1-17.
2. Chao, S. H., Smare, D. L. and Taylor, W. H. *Miner. Mag.*, 1939, **25,** 338.
3. ——— and Taylor, W. H. *Proc. Roy. Soc., Ser. A.*, 1940, **174,** 57.
4. Ito, T. .. *X-Ray Studies on Polymorphism*, Marusen Co., Tokyo, 1952.
5. Mackenzie, W. S. and Smith, J. V. *Amer. Mineral.*, 1955, **40,** 707.
6. ——— .. *Ibid.*, 1955, **40,** 733.
7. Smith, J. V. and Mackenzie, W. S. *Ibid.*, 1956, **41,** 405.
8. ——— .. *Ibid.*, 1958, **43,** 872.
9. ——— .. *Acta Cryst.*, 1954, **7,** 380.
10. Raman, C. V., Jayaraman, A. and Srinivasan, T. K. *Proc. Ind. Acad. Sci.*, 1950, **32 A,** 123.
11. ——— and Jayaraman, A. *Ibid.*, 1953, **37,** 1.
12. Spencer, E. .. *Miner. Mag.*, 1930, **22,** 291.
13. Tuttle, O. F. and Bowen, N. L. *Journ. Geol.*, 1950, **58,** 572.
14. Ito, T. and Sadanaga, R. *Acta Cryst.*, 1952, **5,** 441.
15. Laves, F. .. *Jour. Geol.*, 1952, **60,** 436, 549.

EXPLANATION OF PLATES

PLATE VI

FIG. 1. *b*-axis oscillation photograph of Coimbatore moonstone, oscillated through a range of 15° with the (001) plane in the middle of the oscillation range. Fainter reflections can be seen accompanying many of the spots. The (242) reflection is enclosed in square brackets and the subsidiary reflections can be seen, one on either side of the layer line.

4 JAN 1965

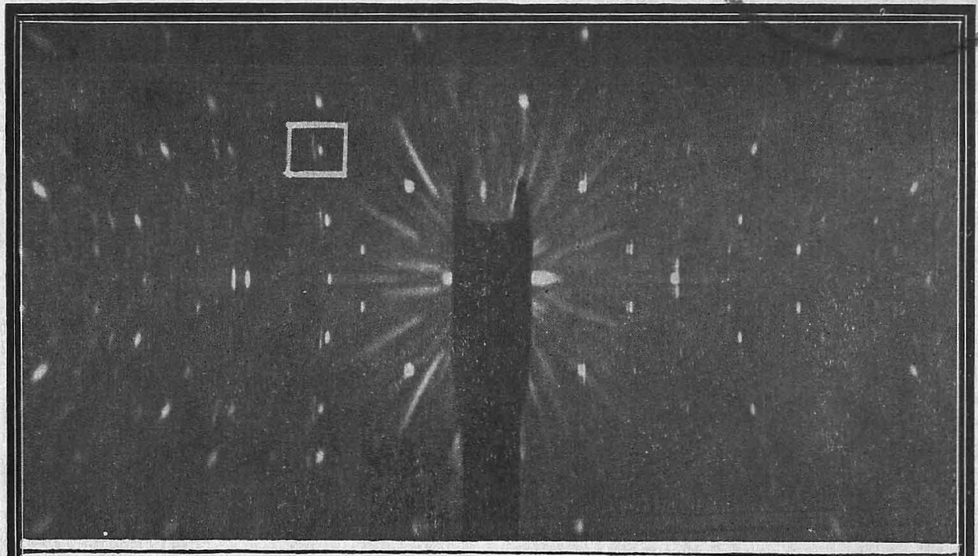


FIG. 1

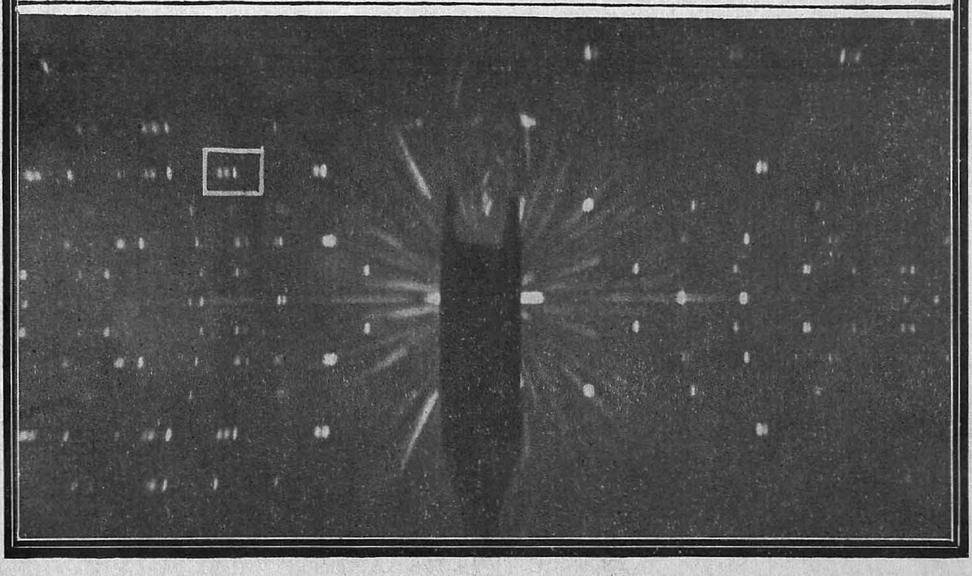


FIG. 2

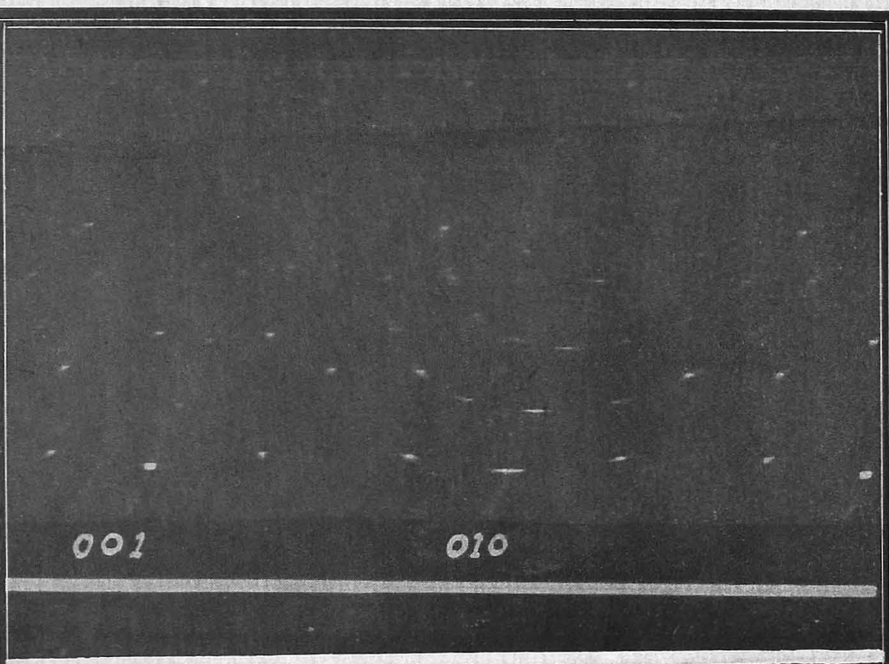


FIG. 3

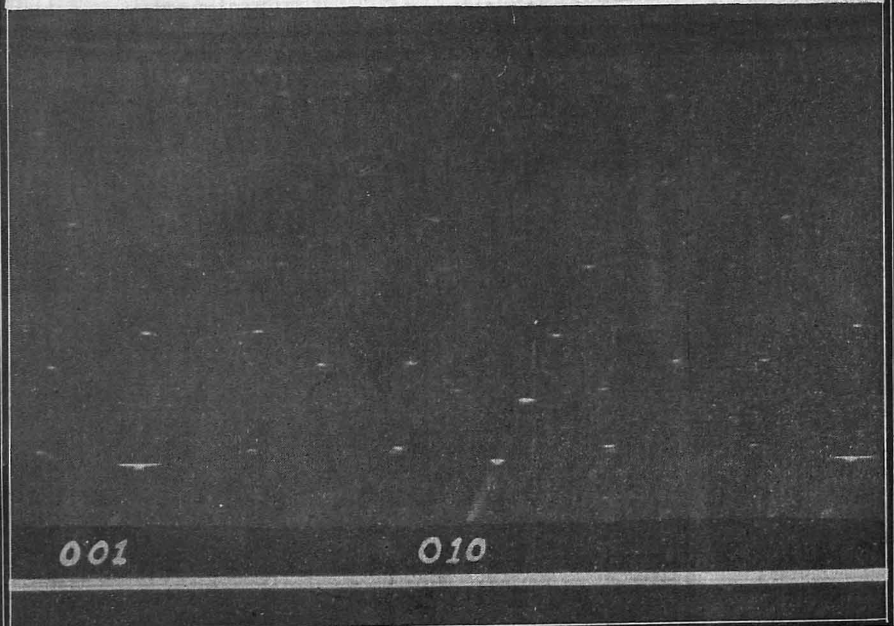


FIG. 4

FIG. 2. Same oscillation photograph as Fig. 1 with Korean moonstone. Principal reflections are accompanied by subsidiary reflection, but the latter fall in the layer line. The (442) reflection is enclosed in square brackets. The two subsidiary reflections can be seen on the left.

PLATE VII

FIG. 3. *a*-axis zero layer Weissenberg photograph of the Korean moonstones. The 001 and 010 reflections are marked. The triple character of the 010 reflections can be seen clearly on high angle spots.

FIG. 4. Same as Fig. 3, but with Coimbatore moonstone. The tripling of the 001 reflections can be clearly seen in the photograph.